DOCUMENT RESUME

ED 325 125 IR 053 321

AUTHOR Shaw, Suzanne J.
TITLE Administration of Library-Owned Computer Files. SPEC Kit 159.
INSTITUTION Association of Research Libraries, Washington, D.C. Office of Management Studies.
REPORT NO ISSN-0160-3582
PUB DATE Dec 89
NOTE 147p.
AVAILABLE FROM Association of Research Libraries, Office of Management Services, 1527 New Hampshire Ave., NW, Washington, DC 20036 ($20.00 per copy; subscriptions are available).
PUB TYPE Tests/Evaluation Instruments (160) -- Reports - Evaluative/Feasibility (142) -- Collected Works - General (020)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Cataloging; *Databases; *Information Management; Information Retrieval; *Information Storage; Instructional Materials; Library Acquisition; *Library Administration; Library Collection Development; *Library Instruction; Library Role; Library Surveys; Longitudinal Studies; Optical Data Disks; *Policy Formation

ABSTRACT

This document reports the results of a follow-up survey of 34 Association of Research Libraries member libraries which was conducted in 1989 to measure changes that had taken place in the administration of computer files (CF)--previously referred to as machine readable data files--since the original survey in 1984. It is noted that this survey purposely took a broader look at the topic, touching on all aspects from selection and acquisition through maintenance and preservation. The introduction begins by making a case for the library as the most appropriate place for storing and making such files available to users. It then summarizes the results of the current survey and briefly discusses several concerns voiced by survey respondents: (1) treatment of CF as any other format when acquiring and cataloging them is inappropriate; (2) the MARC bibliographic format for CF is inadequate; (3) preservation awareness is lacking; (4) the level of service is increasing; and (5) an overall view of CF is lacking. This kit contains the 1989 survey questionnaire and an analysis of the survey data, together with materials submitted by survey respondents. These materials include eight statements of policies and guidelines; two directories; two statements of cataloging policies and guidelines; two course descriptions; and six instructional guides. Institutions submitting materials were Colorado State University, Emory University, Indiana University-Bloomington, New York University, The Ohio State University, Rutgers-The State University of New Jersey, Stanford University, University of Florida, University of Houston, University of North Carolina-Chapel Hill, University of Southern California, and University of Toronto. (16 references) (SD)

SPEC KIT

November / December 1989 OMS

Systems and Procedures Exchange Center

ISSN 0160 3582

U.S. DEPARTMENT OF EDUCATION

ED325125

Administration of Library-Owned Computer Files

Kit 159

IR053321

PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

ARL: OMS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Office of Management Services, Association of Research Libraries
1527 New Hampshire Avenue, N.W., Washington, D.C. 20036 • 202 232 8656

SPEC KIT SUBSCRIPTIONS

SPEC KIT SUBSCRIPTIONS INCLUDE:

- Ten Kits per year (combined issues July-August and November-December)
- With each Kit, five extra copies of the flyer are included for routing to library staff members

NEW SUBSCRIBERS receive a special package with their first issue that contains back issues of OMS publications and Flyers (as available), plus a One-Time Order Form to purchase back issues of Kits for half-price.

- Updated indexes and listings of SPEC materials and OMS publications twice per year
- Advance announcements of OMS training programs and other activities
- The OMS Annual Report

INSTITUTIONAL SPEC FLYER SUBSCRIPTIONS. Many libraries use SPEC FLYER SUBSCRIPTIONS as an inexpensive alerting service for staff members. Flyer subscriptions include:

- Five copies of each Flyer (ten per year)
- Updated indexes and listings of SPEC materials and publications twice per year
- Advance announcements of OMS training programs and other activities

PERSONAL SPEC FLYER SUBSCRIPTIONS. This new service is designed to be an inexpensive way for librarians to keep up on library management issues of today. Personal subscriptions include one copy of each Flyer (ten per year)

TO ORDER SUBSCRIPTIONS

CONTACT YOUR LIBRARY DISTRIBUTOR, SUBSCRIPTION AGENT, OR VENDOR. OR USE THE DIRECT ORDER FORM

OMS

pacflyer
#159 · Nov.-Dec. 1989
OMS
Systems and Procedures Exchange Center
ISSN 0160 3582

ɔMINISTRATION OF LIBRARY-OWNED COMPUTER FILES

By Suzanne J. Shaw
Automation Librarian
University of Florida Libraries

RODUCTION

arch libraries have been evolving beyond being exclusive-epositories of printed matter. Increasingly, the use of erical data files and computer-assisted analysis is becom-the foundation for research efforts in a wide variety of s.

puter files (CF) have traditionally been created or ired and stored by data centers. This was logical, even itable, in an age when computer was synonymous with nframe. Until recently, librarians have been content with rring researchers to such facilities and letting them fend hemselves in using these materials. With the arrival of ocomputers and their ability to remotely access and nload data, and the widespread availability of CD-COM nology with its large storage capacity, this hands-off cy is being reconsidered. A convincing case can be made libraries, not data centers or other remote locations, are proper location for all information.

ere are a number of advantages that make the library the t appropriate agency for the management of CF. Librari-have experience with controlled environments for other s of materials. For example, rare books have long been the eficiaries of appropriate climate control. More recently it been realized that most modern books are endangered to the high acidity level of their paper, prompting onal level preservation efforts. Sensitivity to these concerns naterial preservation finds ready application in computer maintenance.

oraries have established acquisitions and check-in proce-es that are analogous to the updating and record-keeping ired by CF. Many data centers do not have the experience notivation to perform these functions at the level of detail ired for research. Materials in different formats that are integrally related, such as maps accompanying books and various multimedia kits, are also familiar to libraries. Library record keeping procedures make provisions for logically linking the various components.

With a long tradition of assisting and catering to the needs of scholars, librarians have developed a healthy respect for the importance of authenticity. While traditional cataloging methods may not serve to solve all the unique problems presented by CF, they at least provide reasonable assurance that a particular bibliographic entity matches its nominal title, and the record describing it accurately reflects its contents.

Libraries have always acquired resources too expensive for individuals or research entities to provide for themselves. CF of major importance often are just such resources. Since they often cross traditional disciplinary lines, it is possible for several bibliographers to combine their funds in order to purchase CF of mutual interest to their respective fields.

When CF are integrated into the library's finding aids, such as the on-line public access catalog, they become accessible to users seeking information by means of a subject approach. Users who might never think of CF as a resource are made aware of their potential relevance. Moreover, by locating the media within the library rather than just redirecting the patron to other locations, "one-stop shopping" convenience is provided. Reference librarians trained in CF use are available to act as an interface, helping the patron select and use the appropriate data in the optimal medium (which in some cases might turn out to be hard copy). Data center personnel sometimes lack the public service orientation librarians take for granted, and are not prepared to cope with the demands and needs of the public.

The present survey selected the same group of libraries that were surveyed in 1984, both to measure the changes that had

The Systems and Procedures Exchange Center (SPEC) is operated by the Association of Research Libraries, Office of Management Services, 1527 New Hampshire Avenue, N.W., Washington, D.C. 20036. Telephone (202) 232-8656.

taken place, and because these libraries were known to have an interest in the area. The term computer files was used instead of machine readable data files because the Library of Congress has adopted the term, and also because it is more generic, encompassing computer programs as well as data files. It was further decided to limit the scope of the survey to computer files the library owns. This automatically excludes library networks as well as BRS and the like. Also excluded was utility software intended for staff use.

This survey purposely took a broader look at the administration of computer files, touching on all aspects from selection and acquisition through maintenance and preservation. It was felt that this approach could well serve as an educational tool, suggesting to a library that these issues should be considered if they haven't been up until now. These formats require a more elaborate care and feeding than do printed counterparts, and this ongoing commitment should be a part of any strategic planning concerning them.

SURVEY RESULTS

Twenty-eight out of 34 libraries replied to the survey, and of those, 25 acquire computer files, a significant increase since 1984. Nearly half of those surveyed have collection development policies in place that deal with CF, and their selection is most often the responsibility of subject bibliographers. Most have cooperative purchasing agreements with other agencies, usually academic departments. Normal acquisitions procedures are used in the majority of cases. Most libraries provide full AACR2 cataloging for computer files which is made available through online catalogs and the large bibliographic utilities. All feel the need to supplement this information with other finding aids, most often in the form of handouts and flyers.

Surprisingly, a majority of libraries perform no ongoing maintenance after the making of backups upon receipt, and most have controlled only some environmental factors for storing their computer files. One fourth report that the storage environment is completely uncontrolled. There is no standardization as to where maintenance is performed and by whom. In no instance was a preservation unit mentioned as having involvement with computer files.

Access to computer files is typically library use only. Reference staff are the most likely to assist patrons, and the use of staff especially assigned to help with computer file use has risen dramatically. The services most frequently provided are advice on the selection of files and the programs needed to manipulate the data. About half provide data retrieval, and a similar number circulate documentation. The in-house workshop is the predominant form of training for both staff and patrons. Additional training for patrons is widely varied.

ISSUES

Responses to the survey suggest a number of issues to be addressed by libraries concerned with the administration of computer files.

CF are treated as any other format for Tech Services. The effort to treat CF as just another format when acquiring and cataloging them tends to obscure the fact that there are significant differences that directly impact procedures and work flows. For example, selectors and acquisitions staff must consider a number of compatibility factors before the order is placed, as the same data is often available in a wide range of formats. At some point upon receipt, it must be verified that the CF is in fact the one that was ordered, and backups must be made. In order for catalogers to fully describe the CF, hardware on which to mount it must be made available to them.

The MARC bibliographic format for CF is inadequate. The widespread provision of finding aids in addition to the library's online catalog suggests that the MARC format for computer files is inadequate for describing them in the detail their users require. There are problems involved with the multiplicity of versions (expansions, corrections, extractions, reformatting, etc.) that are not amenable to the traditional concept of "editions." Researchers often need information at the record and field level in order to identify sources appropriate for their purposes. This mandates either the presence of a trained librarian as an intermediary, or more detailed information systems in either print or online form.

Preservation awareness is lacking. This survey suggests that the present increase in preservation awareness has not extended to computer files. The rate at which magnetic media can deteriorate is dizzying when compared to that of even the most acidic paper. It is recommended that tapes be cleaned and rewound at least once a year, and that diskette media be mounted and verified at similar frequency. Those found to be unreadable should be restored from backup. Data on magnetic media should be rewritten at least every three years. These intervals assume optimal environmental conditions. Optical media cannot at this point be backed up or rewritten They have not been in existence long enough for their usable lifetime to be known, and they may not be as indestructible as was once thought. At the very least, they profit from a dust- and dirt-free environment.

Level of service is increasing. At the time of the 1984 survey, the most common form of service offered was retrieval of data from online vendors like BRS and Dialog. The next most common service was referral outside the library. This survey indicates that specially trained staff now offer a wide variety of services including the circulation of documentation and in some cases actual programming services to their users.

Overall view of CF is lacking. Very few libraries responding to the survey displayed a broad-based understanding of all aspects of CF management. The absence of this overview hinders strategic planning and threatens academic research libraries' efforts to remain at the center of emerging campus information networks. This fragmented vision of CF is also likely to result in ad hoc, over-simplified training program which work at cross-purposes to each other. What is needed is consensus on basic terminology and principles of operations to guide libraries past the problems presented by CF and help realize the potential locked in these formats.

SPEC Flyer Copyright© 1989 by the Association of Research Libraries, Office of Management Services. This SPEC Flyer may not be reproduced in any form without permission from the publisher.

SPEC Kits (ISSN 0160 3582) are available mainly by subscription. Individual back issues may be ordered for $30.00 ($20.00 for ARL members), with prepayment required, from: SPEC, Office of Management Services, 1527 New Hampshire Ave., NW, Washington, DC 20036. Send check payable to "ARL Office of Management Services." Add $5.00 if you wish first-class postage and handling within the U.S. For all orders to Canada, add $5.00 for postage and handling; for all orders to other countries outside the US, add $8.50 for postage and handling. Kits are shipped library-rate unless otherwise specified; allow four to six weeks for delivery.

spec

Table of Contents

OMS

Systems and Procedures Exchange Center

ISSN 0160 3582

Administration of Library-Owned Computer Files
Kit #159
October 1989

SURVEY RESULTS

Survey Instrument 5
Survey Data Analysis 9

POLICIES AND GUIDELINES

Colorado State University
CD-ROM Task Force Recommendations 15

Emory University
CD-ROM Procedures 19

University of Florida
Criteria for MRDF Acceptance 23
MRDF Decision Form 24
Machine Readable Datafiles at the University of Florida and the Cataloging of Datasets Not Held by the Library 25

Indiana University-Bloomington
Indiana University Libraries-Bloomington and Information Technology 33
Online Database Forum Report 34

New York University
Collection Development Policy Statement: Computer-based Information Sources 41

University of North Carolina-Chapel Hill
The Machine Readable Data Files Center 45
Collection Development Policy for Machine Readable Data Files on Disks 49

The Ohio State University
CD-ROM Technology Policy Recommendations 55

Rutgers-The State University of New Jersey
Checkout Policy for Computer Software 57
Software Information Center-Faculty Borrower Form 58

CATALOGING POLICIES AND GUIDELINES

University of Florida
Cataloging MRDF at UF-A Revised Policy Proposal 61

Rutgers-The State University of New Jersey
Software Cataloging Project 67

DIRECTORIES

University of Houston
CD-ROM Databases 77

Rutgers-The State University of New Jersey
Software Available for Evaluation at CCIS 79

COURSES

New York University
Covering the Bases: Techniques for Data Research and Retrieval 89

University of Florida
MRDF Orientation Course 90

INSTRUCTIONAL GUIDES

University of Southern California
USCInfo Periodical Indexes 93

University of Houston
ABI/INFORM (Flip chart) 113
ABI/INFORM Strategy Worksheet 123

New York University
Self-Service Computerized Information Retrieval 125

The Ohio State University
Guide to the US Census Data Center 129

Stanford University
Academic Data Service Users' Guide 131

University of Toronto
DATALIB Manual 145

SELECTED READINGS

Selected Resources 153

SURVEY RESULTS

March 1, 1989 OMS
Systems and Procedures Exchange Center

The Administration of Library-Owned Computer Files

Your institution has been selected to participate in gathering information and collecting documents for a SPEC Kit/Flyer scheduled for publication in mid-1989. The new kit will update and refocus the information presented in Kit 105, published in 1984. Your library was selected because you participated in the previous survey, or because you have expressed interest in this area.

This survey is designed to be completed by the person responsible for computer files in your library. Responses, **plus accompanying documents,** are due by **April 3, 1989** to:

Suzanne J. Shaw
Automation Librarian
Systems Office, Library West
University of Florida Libraries
Gainesville, FL 32611

If you have questions, please call Suzanne J. Shaw at (904) 392-0796 (Bitnet CYZEU31@NERVM)

NAME AND TITLE OF RESPONDENT _____

LIBRARY/UNIVERSITY _____

PHONE NUMBER (_____)_____

E-MAIL _____

Introduction

We are interested in learning about practices pertaining to the selection, acquisition, cataloging, processing, access , maintenance, and preservation of computer files owned by the library. Ownership is defined here to mean materials purchased entirely or in part with library funds or which have been donated to the library. Included are computer files on all media (e.g., discs, tapes, CD ROMs, etc.). NOT of interest are services provided by on-line vendors, bibliographic databases, or utility software intended for staff use. When response is "Other," feel free to provide comments on the back of the page.

1. Does your library acquire computer files (CF's) as defined above?

☐ Yes ☐ No (If NO, indicate and return questionnaire)

Collection Development

2. Does your collection development policy address CF's? ☐ Yes ☐ No
(If YES, attach policies, guidelines and/or budgets.)

3. Indicate agencies with whom you do cooperative CF purchasing. (All applicable)

- ☐ None
- ☐ Academic Department(s)
- ☐ Computer Center
- ☐ Research Unit(s) or Center(s)
- ☐ Other libraries or networks
- ☐ Other _____

4. Indicate who has authority in selecting CF's. (All applicable)

- ☐ Unknown
- ☐ Subject bibliographers or specialists
- ☐ Format bibliographers or specialists
- ☐ Faculty members
- ☐ Other _____

Acquisitions

5. Are CF's handled as normal materials acquisitions? ☐ Yes ☐ No

6. How are serial CF's checked in?

☐ Not checked in ☐ Normal procedures ☐ Other _____

7. Is CF on-order information available in an on-line system? ☐ Yes ☐ No

Cataloging

8. Is information about CF's available in an on-line catalog? ☐ Yes ☐ No

9. What level of cataloging is provided for CF's? (All applicable)

- ☐ None
- ☐ Full level AACR2
- ☐ Modified AACR2
- ☐ Minimal level AACR2
- ☐ Classification
- ☐ Other _____

10. Who creates cataloging records? (All applicable)

- ☐ Not created
- ☐ Format cataloger
- ☐ Subject cataloger
- ☐ Person outside Catalog Dept.

11. Indicate national database to which CF records are sent. (All applicable)

- ☐ Not sent
- ☐ OCLC
- ☐ UTLAS
- ☐ RLIN
- ☐ Other _____

12. Indicate forms of access other than on-line catalog. (All applicable)

- ☐ Indexes
- ☐ Directories
- ☐ Finding aids
- ☐ Handouts and flyers
- ☐ Other_____

ATTACH EXAMPLES

Processing

13. Indicate places where CF's are processed and/or stored. (All applicable)

- ☐ Systems Office
- ☐ Special Collection or Library
- ☐ Reference
- ☐ General Collection
- ☐ Computing Center
- ☐ Other _____

Maintenance and Preservation

14. Indicate maintenance procedures performed on receipt of CF's. (All applicable; attach procedures)

- ☐ None
- ☐ Backups of selected CF's
- ☐ Backups of all CF's
- ☐ Data verification
- ☐ Other _____

15. Indicate ongoing maintenance procedures. (All applicable; attach procedures)

- ☐ None
- ☐ Cleaning
- ☐ Recopying
- ☐ Recovery from backup
- ☐ Data verification
- ☐ Other_____

16. Indicate who performs maintenance tasks. (All applicable; attach job descriptions)

- ☐ Not Performed
- ☐ Special Collection or Library
- ☐ Systems Office
- ☐ Computer Center
- ☐ Preservation Office
- ☐ Other_____

17. Are climate and other environmental factors controlled? (Attach preservation documents)

☐ All ☐ Some ☐ None

Access

18. Indicate all ways patrons have access to CF's. (All applicable; attach policies)

- ☐ Use in library only
- ☐ General circulation
- ☐ Telecommunication
- ☐ Use only in Computer Center
- ☐ Other_____

19. Indicate staff that assist directly with access to CF's. (All applicable)

- ☐ Reference staff
- ☐ Circulation staff
- ☐ Systems staff
- ☐ Reserve Desk staff
- ☐ Special Collection staff
- ☐ Computer Center staff
- ☐ Other_____

20. What services are provided? (All applicable; attach descriptive information)

- ☐ None
- ☐ Advice on selection of CF's
- ☐ Circulation of documentation
- ☐ Advice on using programs
- ☐ Library staff retrieval of data
- ☐ Programming services by library staff
- ☐ Library staff assigned to help CF users
- ☐ Other_____

Training

21. What training is given to staff? (All applicable; attach descriptive information)

- ☐ None
- ☐ Library-conducted workshops
- ☐ Workshops by outside experts
- ☐ CAI/Programmed instruction
- ☐ On job training
- ☐ Academic course work
- ☐ Other_____

22. What training is offered to patrons? (All applicable; attach descriptive information)

- ☐ None
- ☐ Library-conducted workshops
- ☐ Workshops by outside experts
- ☐ CAI/Programmed instruction
- ☐ Other_____

BE SURE TO INCLUDE **DOCUMENTS** YOUR LIBRARY CAN SHARE ON LIBRARY OWNED COMPUTER FILES. RETURN SURVEY AND DOCUMENTS BY **APRIL 3, 1989** TO:

Suzanne J. Shaw, Automation Librarian
Systems Office, Library West
University of Florida
Gainesville, FL 32611
(904) 392-0796

34 libs

12

Survey Data Analysis

Data from the survey was input into a dataset and analyzed using SAS in the CMS environment. The numbers below correspond to the questions on the survey.

1. This tabulation considers only those who acquire CF (25 of 28 respondents).

COLLECTION DEVELOPMENT

2. Collection Development (CD) policies address CF in 44% of cases.

3. There are no cooperative purchasing agreements in 44% of cases. Of the remaining 56% who do have them:
   44% have them with academic departments.
   24% have them with a computing center.
   12% have them with research units.
   24% have them with other libraries or networks.
   There were no other agreements mentioned.

4. 4% did not know who had CD authority.
   88% report subject bibliographers/selectors have CD authority.
   20% report that format bibliographers/selectors have it.
   28% report that faculty have it.
   40% report that authority lies with others:
   - 12% library administration
   - 12% data librarian/archivist
   - 8% reference staff
   - 4% systems staff
   - 4% received on deposit

ACQUISITIONS

5. 75% of respondents handle CF as normal materials acquisitions.

6. 20% do not check-in serial CF.
   58% check them in using normal check-n procedures.
   20% handle serial CF in some other way:
   - 4% checked in by owning location
   - 4% checked in by systems
   - 4% code books only checked in
   - 4% title records, no item check in
   - 4% procedures vary

7. 50% indicate that on-order information on CF is available on-line.

CATALOGING

8. 68% of respondents report that bibliographic information about CF is available online.

9. 28% of respondents report that no cataloging is provided for CF.
72% provide full AACR2 cataloging.
None provide modified AACR2 cataloging.
8% provide minimal AACR2 cataloging.
12% classify CF.
16% provide some other level of cataloging access:
    4% database on mainframe
    4% brief manual catalog
    8% unspecified

10. 20% do not prepare cataloging records.
58% report they are prepared by format catalogers.
25% are prepared by subject catalogers.
20% say they are prepared by someone outside catalog dept.:
    4% data librarian
    4% computer center staff
    12% unspecified

11. 24% do not send records for CF to a national database.
56% send them to OCLC.
4% send them to UTLAS.
48% send them to RLIN.
4% send them elsewhere (National Union Catalog).

12. 12% provide additional access through indexes.
16% provide it through directories.
20% provide it through finding aids.
64% provide it through handouts or flyers.
28% provide other forms of additional access.
    8% card catalog
    8% database on mainframe
    4% cataloged code books
    4% word of mouth
    4% unspecified

PROCESSING

13. In 28% of cases, CF are processed by and stored in systems.
28% are in special collections.
64% are in reference.
28% are in the general collection.
52% are in a computing center.

28% are handled elsewhere:
- 8% technical services
- 8% reserve room
- 8% academic departments
- 4% branch library

MAINTENANCE

14. In 36% of cases, no maintenance is done on receipt of CF.
32% make selected back-ups.
40% make back-ups on all receipts.
28% do some data verification.
No other maintenance on receipt was reported.

15. 56% perform no on-going maintenance of CF.
20% perform on-going cleaning.
36% perform on-going recopying.
36% perform recovery from back-ups.
12% perform on-going data verification.
8% perform some other kind of on-going maintenance:
- 4% spin tapes annually
- 4% unknown

16. In 32% of cases, no one performs maintenance.
12% report maintenance is performed by special collections staff.
36% report it is done by systems staff.
28% report it is done by computing center staff.
No one reported it done by preservation staff.
20% report it done by other staff:
- 4% data archivist
- 4% faculty members
- 4% reference staff
- 4% holding units
- 4% under review

17. 24% report no environmental factors are controlled for CF.
68% report some are controlled.
8% report all are controlled.

ACCESS

18. 80% allow library use only of some CF.
28% permit general circulation.
40% provide access by telecommunication.
36% allow computer center use only.
20% provide some other mode of access:
- 8% campus network
- 4% subsets on diskette

4% faculty
4% unspecified

19. 88% have reference staff assist patrons with CF.
4% have circulation staff do it.
36% have systems staff do it.
32% have reserve staff do it.
20% have special collections staff do it.
40% have computer center staff do it.
16% have other staff do it:
    8% faculty
    4% data archivist
    4% government publications staff

20. 4% reported no services are offered to patrons.
80% provide selection advice.
48% circulate CF documentation.
80% provide program advice.
52% perform data retrieval.
20% offer programming services.
72% have staff especially assigned to help with CF.
20% provide some other sort of service:
    4% guides
    4% referral to computer center
    4% open houses for CD Products
    4% online documentation
    4% local documentation

TRAINING

21. 8% provide no staff training.
64% provide library workshops for staff.
44% utilize outside workshops for staff.
None use CAI/PI for staff.
84% provide on-job training.
12% use academic course work.
4% provide some other sort of staff training (documentation from vendor/agency).

22. 16% provide no patron training.
60% provide library workshops for patrons.
20% utilize outside workshops for patrons.
8% use CAI/PI for patrons.
48% provide some other sort of patron training:
    28% one-one-one consulting
    8% presentation to academic classes
    4% overview open-house programs
    4% on-site instruction booklets
    4% integration with existing BI

POLICIES AND GUIDELINES

COLORADO STATE UNIVERSITY LIBRARIES

OFFICE MEMORANDUM

TO: CD-ROM Task Force

FROM: Nancy Allen

SUBJECT: Recommendations

DATE: November 21, 1988

CD-ROM Task Force

The CD-ROM Task Force, chaired by Nancy Allen, consists of Fred Schmidt, Suzanne Johnson, Nancy Ellison, Jennifer Monath, Judy Berndt, Halcy Enssle, and Barbara Burke.

The Task Force met once, convened and chaired by Fred Schmidt, and then four more times, chaired by Nancy Allen. After these discussion, the Task Force is ready to forward draft recommendations on the use of CD-ROM products. It is the intent of the Task Force to seek very wide input on these recommendations, and on the basis of such input, revise them for final review by the Director. Please return any comments to N. Allen by Dec. 2.

Policy Recommendations

1. We will try not to discriminate between non CSU and CSU-affiliated users unless and until there is difficulty scheduling users. If scheduling conflicts arise, CSU affiliated users should receive priority.

2. Uncontrolled, unscheduled access to the workstations should be allowed when there is no queue and when no user has scheduled use of the workstation in advance. All workstations will also be available on a pre-scheduled basis, and the scheduled patron has priority over others. Scheduled time is reserved in 30 minute blocks, with no more than two such blocks of time reserved per day per person.

3. We will not require that patrons receive formal training or instruction prior to the first use, but we will make such instruction available, either through individual tutorials, brief personal assistance, small-group or classroom instruction, or written instructions. If a CD product is not intended to be end-user oriented for the general user, the librarian will perform the search for the user unless the user intends repeated use of the product, in which case instruction should be provided.

4. If a patron needs more than one disk, the patron should receive instruction in changing disks, or this should be done by the librarian.

5. Each department will maintain and distribute to all other reference desks written information on the contents, coverage, storage and security, and off-hours availability of their disks. Any information items necessary for use of equipment will also be included. CD products usually housed on the second floor will be available during Science/Technology reference hours. Access at other times will be provided as staff availability permits.

6. We will not limit printing unless paper costs become too great, or unless it is apparent that a user is abusing the print function. Users should be encouraged to download bibliographies to floppy disks.

7. The microcomputer services librarian, in cooperation with the appropriate librarian from each department, will be responsible for training all desk staff in the use of all disks, including downloading to floppy disks.

8. Librarians will be responsible for end user instruction in their assigned subjects. If a product is interdisciplinary or non-disciplinary, open workshops on the product should be offered on a regular basis.

9. General information on CD products will be included in relevant guides, handouts, and CO 150 materials. Librarians are strongly encouraged to include information on these products in relevant special bibliographies or guides prepared for classroom instruction.

10. Signage should be used to identify workstations dedicated to a single data file.

11. Librarians should not do mediated searches for a fee when the same results could be gained through use of a CD product without explaining all options to the patron. Librarians, assisted by the Microcomputer Services Librarian, should be aware of costs and benefits for CD and for mediated, telecommunication-based searching. CD-ROM products are intended for direct patron access. Mediated search requests will be handled through CLASS, with few exceptions (as for services for patrons with physical disabilities).

12. Selectors should consider factors enumerated in the CD-ROM Evaluation Checklist which appeared in the May 1988 issues of Laserdisk Professional in making selection decisions.

13. Funding for new CD products should be considered in the same way as new reference serials. New titles should be justified and reviewed by the departmental serial committees of the whole, and funding for subscriptions should be allocated in coordination with Joel Rutstein as Collection Development Officer. It is the selector's responsibility, assisted by the Department Head and the Head of Public Services, to plan for acquisition of the necessary equipment for workstations.

14. Whenever possible, new CD purchase costs should be offset by cancellation of duplicated print or microform resources. Space concerns should also be taken into account in making these decisions.

15. It is the selector's responsibility to determine copyright restrictions for downloading.

lw

EMORY UNIVERSITY GENERAL LIBRARIES
TECHNICAL SERVICES AND COLLECTION MANAGEMENT DIVISIONS

CD-ROM PROCEDURES

Collection Management:

Selectors should send any order for a CD-ROM subscription, whether for a trial period or firm commitment, through Collection Management. Collection Management will then send the Acquisitions Dept. typed order slips with any corroborative information attached. The specific library department destination should be indicated on the bottom of each order.

Examples:

AUTHOR (Last Name First)
General Science Index. (WilsonDisc CD-ROM). To be

TITLE
updated quarterly.

TITLE (continued) | ISBN/ISSN

Place of Publication	Publisher	Date	Price
NY	H.W.Wilson	1988-	$1,295.00

SERIES:

DEALER: CAT. NO.: ITEM NO.:

REQUESTED BY: BG DEPT: SCI REF DATE: 9/21/88
New Subscription Cat for Sci Ref

AUTHOR (Last Name First)			
General Science Index. (WilsonDisc CD-ROM). To be			
TITLE			
updated quarterly.			
TITLE (continued)		ISBN/ISSN	
NY	H.W.Wilson	1988-	$1,295.00
Place of Publication	Publisher	Date	Price
SERIES:			

DEALER: CAT. NO.: ITEM NO.:

REQUESTED BY: BG DEPT: SCI REF DATE: 9/21/88

Trial subscription Cat for Sci Ref

ALL CD-ROMS are RUSH items and should be accompanied by the Collection Management RUSH slip.

Acquisitions Department/Serials Control Department

The "ship-to" address will be Acquisitions Dept./Serials. Upon arrival, the initial material will be noted on the goldenrod slip in the Kardex and sent to the Accounting Unit as any other new subscription. Accounting will notify the department where the CD-ROM will be used to pick it up for testing. If the CD-ROM package is satisfactory, it will stay in the "home" department and the Accounting Unit will be notified that it is "o.k. to pay." Accounting will then send the paperwork to the Head of Serials Control for cataloging purposes. However, full cataloging will not take place until a decision has been made to continue the title past the 1 year trial period.

Serials Control will set up a regular Kardex check-in record.

CD-Roms which need to be returned:

1) Add note to Kardex: "Check in and deliver unmarked immediately to [dept.]. Notify [dept.] to return previous update to Serials Control."

2) Serials Control will save the mailing materials received with the new update. When the dept. returns the previous update, Serials Control will put it in the wrapper provided for mailing and take it to the mail room for return to distributor.

22

CD-ROMs which are discarded:

1) Add note to Kardex: "Check in and deliver unmarked immediately to [dept.]. Notify [dept.] to discard previous update."

2) Serials Control will assume all CD-ROMs which do not come with mailing materials to return the previous update may be discarded. (Note: H. W. Wilson CD-ROMs: Business Periodical Index, Readers' Guide, and MLA International Bibliography, may be discarded.)

Acquisitions Department/Catalog Department

If any CD-ROMs are indicated as monographs, the shipping address will be "Acquisitions Dept." and the material will be checked in as a monographic continuation. Accounting will notify the department for testing and, if the quality of the item is satisfactory, will send the paperwork to the head of the Catalog Dept.

The update will be checked in by the Acquisitions Dept. Continuations Associate who will send it to the designated department with the same instructions for returning the previous update as mentioned above.

Cataloging (general):

CD-ROMs will be given 099 call number:

099 CD-ROM ≠a 1 [etc.]

One sequence of numbers will be used for all CD-ROMs.

Locations will be wr, wsr, etc. as appropriate.

049 will be EMU2, EMUR, etc. as appropriate.

949 ≠t will be "n" (non-circulating).

949 ≠n will not be used for serial CD-ROMs and discs will not be marked or barcoded since no serial disc is a permanent part of the collection. Procedures will be developed as appropriate for any monographic CD-ROMs.

949 ≠z will be used to explain return procedures. Kardex will also be noted. Detailed manual procedures will be developed.

Collection Management/Acquisitions Department/ Cataloging Departments

Collection Management must notify the Acquisitions Department in writing as soon as the decision to continue or discontinue the subscription is made.

If the title is to be continued beyond the trial period, Collection Management will generate a new order to send to the Acquisitions Department. This order should include any changes in specifications and should have the blue and pink slips attached. When the Acquisitions Department receives the order, the blue slip will be given to the Accounting Unit and the pink slip to Serials Control or Catalog Dept. with "Renewal" stamped or written on each slip. Serials Control or Catalog will then do full cataloging on the CD-ROMs. The white slip will be filed in the Current Orders drawer. If the title is to be discontinued, that information will be relayed to the Accounting Unit and Serials Control by the Acquisitions Dept. upon receipt of notification from Collection Management.

Note: If a renewal notice for a title is received by the Accounting Unit from the vendor before the decision has been made to continue the subscription beyond the trial period, Accounting will notify Collection Management, who will make the decision and report back to Accounting.

Criteria for MRDF Acceptance

1. Media/Data Integrity. The media containing the MRDF must be physically undamaged, which is to say the R in MRDF is a requirement. Damage can sometimes be determined by physical inspection of the media, but usually is only detected when we try to make a backup copy and/or map.

2. Adequate Paper Documentation. Unlike most multi-part materials stored in a Library, a MRDF is completely worthless without some sort of documentation which describes its internal format and meaning. The condition of the MRDF itself is irrelevant if such documentation is unavailable. Having the documentation stored on the MRDF medium itself is unacceptable as the sole source of documentation for two reasons. First, we cannot be sure of the integrity of the documentation itself without an independent source against which to check it. Second, we cannot judge the amount of space and processing the MRDF will require without a printed copy of the documentation, and printing a large file can be a problem in itself. The information supplied must also permit the item to be cataloged (if we decide to require cataloging for all MRDF) under whatever cataloging rules we end up adopting.

3. Data "Usability." The MRDF must be suitable for use within the UF environment. Use of the MRDF must not require equipment which is not available for use at NERDC or other designated use sites, nor must it require any special software for processing which is not also supplied with the dataset or otherwise available in the Library.

4. Broad Applicability. The MRDF must be of interest to a sufficiently broad segment of the research community to justify its care and feeding. Although this is especially important for datasets which we purchase, it also applies to datasets which are offered to us at no charge. Resources for preserving and using MRDF are finite, and we do not wish to become a dumping ground for data of interest to only one individual, any more than we would buy an expensive book collection which is likely to be used only once.

5. Feasible Security/Usage Requirements. The provider of the MRDF must not make infeasible requirements for its use by our patrons. We cannot guarantee that the usage of a dataset will be entirely confined to our patrons for research use in many cases, and a dataset which was offered with such a stipulation would be rejected.

6. Feasible Retention Requirements. We cannot promise to keep a dataset forever. Any MRDF we accept must be understood to be subject to a change in status and removal from the archive at a future date. We cannot promise to notify the entire universe of potentially interested individuals of upcoming deletions [We might want to suggest the formation of a clearinghouse for such announcements, however, to be conducted at a national level].

University of Florida Libraries
MRDF Decision Form--Draft

Selector: _____ Systems Office Reviewer: _____

Date: _____ Library Fund(s) Charged: _____

MRDF Title: _____

Version/Level: _____ Frequency of updates expected: _____

Publisher: _____

Vendor address: _____
_____
_____

Anticipated Cost: $_____ Catalog Level Desired: _____

Format Acquired: _____

Equipment Required: _____

Application Software Required: _____

Location for User : _____

Circulation Policy: _____ Copies Made In: _____

Description of User Population (staff, faculty, etc.) and Primary Academic Programs Served.

_____

_____

Description of Local Programming Anticipated: _____

_____

_____

Documentation [ ] Number of Copies: ___ User Copy Location(s): _____

Comments on Documentation: _____

_____

_____

User Restrictions: _____

_____

Approvals: _____

Machine Readable Data Files at the University of Florida and the Cataloging of Datasets Not Held by the Library

By: R. Max Willocks & William C. Covey, III

I. General MRDF Issues

The term Machine Readable Data Files is not yet widely understood. Many faculty members do not understand the term and many of those who do understand it, do not understand why the library is involved. The library position on the other hand is that it has an obligation as the provider of research information resources to the university to include all types of information in all types of formats.[1]

Historically the library has endeavored to train researchers in the use of new media (cf. microforms) and to house them in the library. It is appropriate then for the library to assume that same position with MRDF.

When a library is newly considering getting into the field of managing and servicing electronic data or even if it has been involved for several years, it will be well to ponder and to plan to answer certain questions. The following list, although not exhaustive, is perhaps helpful.

1. Who in the library should be involved?
2. What part should the Bibliographers play?
3. Will MRDF be stored in the library?
   - If so, what will be the access?
   - What special equipment will be needed?
   - What preservation procedures need to be implemented?
4. What help should the library provide to users?
   - Programming?
   - Statistical consultations?
   - Actually running the programs on the mainframe?
5. Should the library consider handling data in all machine formats?
6. If datasets are maintained in the library for use on microcomputers, should the library also provide microcomputers on which to manipulate the data?
7. How will selection for purchase be made?
   - Deselection?
8. How will purchase be budgeted?

---

[1] Other libraries are now accepting this concept. The entire article by Siman is good but one paragraph is especially: "Since the library staff did both bibliographic and numeric searching, they became more acutely aware of the information potential of machine-readable data files. Searching for numeric data in turn also has brought them into closer contact with patrons anxious to use those data for analyses not possible through online access, with the result that more librarians have a sense of the promise of machine-readable data files. These factors, coupled with the diffusion of automation into all the services of the library, have created a current environment much more conducive to the creation of data libraries in academic institutions." Siman, Bliss Beckman, "Machine-Readable Data files in the Library", Technicalities vol. 6, number 6, June 1986.

9. Will the library accept, process and service MRDF gifts or items placed on deposit?
   If so, under what conditions?
10. How will the library best make its collections of MRDF known to and available to prospective users?
11. Where should the user interface occur?
12. Should local MRDF holdings be added to a national data base?
    Will the locally held MRDF be available on inter-library loan?
13. What level of cataloging is appropriate?
    Is staff available to catalog the material?
14. If you decide to catalog MRDF should you consider cataloging MRDF available on campus, but not held by the library?

Another perspective on challenges to be considered can be found in the article by Kwan, Duncan, et al., "Textual and Numeric Data Files".[1] Recommendations which they make are as follows:

RECOMMENDATIONS

1. Needs assessment studies to be conducted
   -- campus libraries work closely with faculty, computing centers, data archives, and other appropriate campus departments to determine local needs, both in the present and with projections for the future.
2. Campus advisory committees be established to advise on local acquisitions and services
   -- should be involved in the needs assessments mentioned above and in developing policy, acquiring budget support, addressing service-related issues, and serving as advocates for campus and university administrations.
3. Responsibility for textual and numeric data files be assigned to specific personnel
   -- coordinate selection of data files and provide consulting and advisory services on the use of these files.
   -- promote awareness and use of data files.
   -- keep abreast of developments in this field.
   -- serve as resource persons on matters related to data files.
   -- serve as liaisons between the library and other departments on campus.
4. Library serves as a central repository of information about textual and numeric data files and specifically, should collect:
   -- published directories of data files and guides to the holdings of data archives which identify the existence and availability of these resources.
   -- core collections of periodicals and monographs devoted to the utilization of these materials and to this developing field.
   -- code books, system manuals, other documentation, if appropriate for campus needs.
5. Library serves as an information referral center
   -- provides cataloging access to data files and related materials following AACR2 and MARC format.
   -- provides referral access for materials or services which do not fall within the above category, either through campus-wide directories of services or through linking records in local online systems.
   -- plans staff development to promote awareness of textual and numeric data files.

---

[1] Kwan, Julie; Duncan, Joyce and others, "Textual and Numeric Data Files", Librarian's Association of the University of California, Berkeley, Oct. 25, 1983.

-- provides reference and referral services relating to data files.

6. The library cooperates on a systemwide level regarding MRDF
   -- resources shall be appropriately shared throughout the campuses and a coordinated plan be developed to provide the best services possible.
7. Contracts be negotiated with provisions for the broadest possible eligible user populations
   -- campus and systemwide participation in contracts be developed, and cooperative sharing programs recommended.
8. Funds be allocated to provide access to data files
   -- campus and systemwide administrations must realize the value of data files for the research and instructional programs of the university and allocate funds to support appropriate acquisitions and services.
9. These resources must be included in long-range planning."

II. How the University of Florida is Addressing Some of These MRDF Issues

The University of Florida decided many years ago to collect MRDF just as it does other formats. Originally these files were selected and purchased by an assistant director. Later a committee was formed for that purpose. The committee was composed of two members of the Systems Department, two members of the Reference Department, one member each from Acquisitions, Cataloging, and Documents and an assistant director. The committee approach was very successful. However, when the library began to implement a collection management program it became more appropriate for the bibliographers and selectors to incorporate MRDF into their regular collection management program. This program is now beginning to function. The bibliographers and selectors are currently being trained through a one month seminar that meets for one hour, three days per week for the month. We believe that they will be much more nearly ready to handle this issue upon completion of the course. The outline of the course being taught by the head of our Systems Department is as follows:

Course Objectives:

The purpose of this short course is to present a rudimentary overview of some of the concepts and techniques needed to work with machine readable data files, or MRDF. Participants will learn to evaluate MRDF for possible inclusion in the Libraries collections on the basis of documentation, suitability for our computer environment, and quality. In addition, participants will develop a better appreciation for the processes by which a dataset can be prepared to describe phenomena. Basic terminology relevant to discussing MRDF in an academic environment will be learned.

Organizational Information:

The course will be conducted for four weeks, from 10 May 88 to 2 June 88. Meetings will be held from 1:00 to 1:50 every Tuesday, Wednesday, and Thursday. The instructor will be Bill Covey (Systems, Library West, 392-0796).

Work outside the class will consist primarily of construction and manipulation of two datasets, one constructed by each participant individually, and one supplied from the Libraries MRDF archive. Some small readings will also be supplied.

Periodically, evaluations will be made of participant progress.

10-12 May

A. Overview, course organization, Data Matrix I
B. Measurement taxonomy, useful arithmetic
C. Arithmetic bases, coding schemes

17-19 May

A. Data organization, media types
B. Documentation, Data Matrix II
C. Sign-on, Data Entry and Editing

24-26 May

A. SAS I - Basic data access and manipulation
B. Data Cleaning I - Frequencies & Crosstabulation
C. Statistics - Measures of central tendency (mean, median, mode)

31 May - 2 June

A. Generic Computer Systems
B. Local Computer Systems
C. Summary

Prerequisites:

Scientific calculator (at least one memory, square roots, logs), Faculty Support Center sign-on, willingness to suspend disbelief.

MRDF are stored in the library under special temperature and humidity controlled conditions. A tape cleaning machine was purchased and the tapes are cleaned on a regularly scheduled basis.

The interface with the user is initiated at the reference desk and may involve a very complex discussion before the researcher is referred on to the Systems Department to make the arrangements for the actual use of the tapes. The library provides limited programming help. Statistical help and sophisticated programming help is available to users from another campus agency. The library delivers the tapes to the University Computing Center for the use of the researcher but does not run the programs on the mainframe for the user except on a limited contractual basis.

The library decided early in the program to collect MRDF in all formats and this has not presented a problem. However, collecting data on floppy disks for microcomputers has produced requests for access to machines on which to use the disks. As of this time the library has not provided such machines because of the lack of space in which to put such machines and the lack of funds with which to purchase them.

Selection and purchase of MRDF are now made through regular collection management and acquisitions procedures. The budget comes from the regular appropriation to the given discipline although several bibliographers may combine their resources in order to make a given purchase.

The library has received numerous gifts of datasets in the past, but is now building a program through which the library will actively seek to obtain datasets held by the faculty.

When these are received by the library they will be processed similarly to other datasets. The library has chosen not to catalog datasets that are not in its possession (discussed below). We are now cataloging datasets held in the library according to nationally recognized standards and we are putting these records in the national data bases. It has not yet been decided how to deal with interlibrary requests for these materials although one major dataset was given to the library on condition that it be available to researchers from other universities.

The interface with researchers has always been through the reference librarians but we are now seeking to involve all the bibliographers in this process and tools are being prepared to enable the library to publicize its MRDF holdings and services.

III. Cataloging of Non-Library MRDF Materials

The issue of whether a research library should catalog MRDF materials which are not held in its collection must be considered in terms of the rationale for dealing with MRDF to any extent. What advantages are there to having involvement of the library for the MRDF user?

We see at least five major advantages which argue in favor of including MRDF in the scope of library activity:

1) libraries are the traditional place for purchase and sharing information resources;
2) libraries provide enhanced access to research resources, in that they have controlled indexing and descriptions, provide a centralized location for access, and maintain systems for circulation and retrieval of materials;
3) libraries provide a reference service interface between the researcher and library materials;
4) libraries engage in preservation efforts for their materials; and
5) libraries provide a measure of authenticity, in that they provide reasonable assurance that a particular work matches its nominal title.

In terms of these hoped-for benefits, it is not clear what is to be gained by cataloging materials which are not under library control. The convenience for the patron of "one stop shopping" for locating MRDF resources of an institution is severely mitigated by cataloging outside datasets. After finding the dataset entry in the catalog, the patron must then try to locate its current owner, somewhere in the service area of the library. For a research library, this owner will frequently be an academic who is not prepared to cope with performing a circulation function for the public. Patrons looking for datasets which are not held by the Library will be forced to adapt to the hours and conditions of contact set by the holder.

If and when the dataset is located, it may not bear much resemblance to its entry in the catalog. The catalog entry can only record the description of the dataset at the time at which it became known to the library. It is unrealistic to assume that an outside holder of a dataset will notify the library of every change in the dataset contents, revision status, media type, and other pertinent variables. It is likely that the catalog entry will become increasingly inaccurate over the useful life of the dataset.

The accuracy of the dataset itself may also suffer two potential erosive effects. The first such problem involves conservation. Individuals are frequently not in a position to provide the environment needed to preserve MRDF. For optimum data life, humidity and temperature must be kept within fairly narrow limits. As magnetic media age, they begin to break down. The process can be slowed by cleaning (on an expensive machine) and proper physical storage (in specially constructed racks), but the only ultimate conservation measure possible is the creation of multiple backup copies. The documentation required to use the dataset must also be

kept in good condition and updated as necessary. These measures must be maintained without lapse for as long as the dataset is kept.

The second accuracy problem involves authenticity. In the process of dataset analysis, researchers frequently modify the original data into new, synthesized forms. Multiple copies of the original dataset proliferate, each altered to some extent from the original source. This may not be a problem for the original researcher, who has gleaned whatever information needed at each stage of the evolution of the dataset, but it is devastating for someone performing secondary analysis. If the original form of the dataset cannot be preserved, confirmatory analysis is impossible. Even when the dataset used for secondary analysis is supposed to be an exact duplicate of the original, any new or contradictory conclusions derived therefrom are suspect when the dataset has been under the indifferent control of one or more intermediaries. Unlike traditional media, MRDF are virtually indistinguishable under physical inspection; the accuracy of their documentation is all that make it possible to tell one from another.

These problems associated with outside control of the datasets all work against the original advantages of having libraries involved with MRDF. To the extent these problems are present, the contribution of a library to research use of MRDF is diminished. A library cannot be of any assistance in alleviating these problems unless it has control of the datasets. The only real advantages of a library of cataloging or otherwise maintaining finding aids for materials it does not own are 1) to enable an occasional researcher to find some material that might not otherwise have been found; and 2) a reduction in acquisition and preservation costs for the library. While it is true that providing proper storage and handling for MRDF entails considerable expense, someone must fund the operation if the datasets are to survive. Libraries are the traditional institution charged with long-term preservation of research resources. An individual researcher can probably maintain MRDF long enough for his immediate personal use with no special care at all. If libraries do not control datasets to provide accurate and current descriptions, ease of access, authenticity, and preservation, then their major contribution to MRDF-based research will be to acquire datasets for shared use which are too costly for other research entities to provide for themselves.

IV. What of Tomorrow?

It is inevitable that datasets will play a larger and larger part in university research and therefore in library holdings and services. For instance, the college of Business at the University of Florida reports that 75% of all articles published in business journals on business subjects depend on computerized datasets. The number of faculty members already using datasets[1] certainly underscores the need of these materials in the future. Therefore, libraries simply must get into the mainstream of research information in computerized form.

Accepting the premise that MRDF are going to be an increasingly important component of collections at serious research libraries has some implications for library resource allocation. The most obvious of these involve collection management and bibliographic control: acquisition, cataloging, and preservation procedures must be established and/or modified to accommodate MRDF. The realities of limited budgets requires us to provide a way to balance these new demands with existing needs. Much of our effort on this project has been focused on these considerations.

---

[1] The University of Florida survey reveals that 59.2% of those faculty returning the survey forms are using electronic data sets.

Access services for MRDF have not received a comparable amount of resources. Our approach to date has been to accommodate existing circulation practices to current patterns of patron use of MRDF. This has been effective, in that patrons have had the benefits of library preservation and bibliographic skills for MRDF, but it is not likely to meet future demands. Our current situation is analogous to a microform archive which has no microform readers, or an audio archive with no playback equipment. We need to provide a way for patrons to utilize MRDF without depending on outside facilities.

To do this, we must acquire our own computing facilities, which are the MRDF equivalent of the readers and tape machines of other specialized formats. Computing facilities used by a community of scholars have several characteristics. First, they are prohibitively expensive relative to the financial resources of most departmental and even college level units. This expense takes the form of acquisition costs, costs of continuing operation, and space. Second, although the initial outlay for the equipment is high, it may be efficiently shared among multiple patrons, making the net cost per user reasonable *if* it is spread over a sufficiently wide base. Third, these facilities are essential for an ever increasing variety of research in almost all disciplines. Fourth, the complexity of these resources requires an increasing number of researchers to seek professional assistance in their use. Finally, there is a certain amount of computer-assisted research (CAR) which requires exploratory computer use. This is seldom the sort of usage which can be immediately supported by specific grant proposals, yet this "browsing" greatly aids in the production of more directed work.

All of these characteristics of CAR are shared by traditional library materials. There is a great communality of function between research use of computing and research use of libraries. The topic of merging the computing center and the library is much debated, but this proposal does not deal with that topic *per se* . Instead, it is argued that if we agree that the library is the appropriate vehicle for the delivery of MRDF to the research community, it is appropriate for the library to have *a* general purpose computing facility to permit efficient MRDF usage.

The advantages for MRDF usage are obvious: media would "circulate" within the library, from data archive to computer and back, without being physically transported to a remote data center. This would increase both physical security for the datasets and provide for more timely access. Since this would be a more specialized facility than the typical campus data center (which is frequently shared with administrative computing), more extensive and useful statistics on usage of MRDF could be kept. Online storage could then be allocated for MRDF during periods of high usage. Preservation MRDF copying would be facilitated by having the necessary computer equipment in the library. Access to computer resources for the research community could be broadened, as the machine would be dedicated to library use and not in competition with administrative functions. An economy of scale would be obtained in terms of controlled storage space, since an enlargement of the area required for proper MRDF storage would provide space for computer equipment. Librarians are increasingly skilled in computer usage, and being proximate to the MRDF usage area would enable them to share those skills with researchers as the analysis is being done. Librarians would also profit from relationships with the library computer staff, which could provide computer expertise in a library setting.

Although a library might find it difficult to justify an additional outlay for large-scale computer equipment as an aid to MRDF alone, other possible applications within the library might combine with MRDF to make the project feasible. The same telecommunications links which would provide access to datasets for remote users could also be used for a variety of person-to-person communications, such as electronic mail between collection management staff and faculty. A very attractive use for such an in-house computer would be to put a subset of commercial bibliographic services online. One of the major problems with end-user use of commercial services is the high cost of search time, sometimes approaching several dollars per minute. This is intimidating both to the novice searcher and to the sponsoring agency. Mounting typical databases on a machine which is free of such time-clock pressures would provide economical practice time for training both end-users and staff, as well as covering some basic research needs.

Many other uses for such a facility can be offered, from general library automation to serving as a communications node in an interlibrary network. Where a library computer is already in place for these or other purposes, it may be possible to expand its usage for MRDF with little or no further expenditure. The ultimate rationale is that libraries have evolved from being defined as warehouses for printed materials to being the primary access point for information services. If they are to keep and justify this role, they need to control their computing resources as the primary tool of the trade.

In an altogether different direction for the future we may find that technology will lead us along lines not presently anticipated. The course of microcomputers may open up vast new approaches. IBM has invented a 3.5 inch magnetic disc that can store ten gigabytes of information, which may mean that what is now thought of as a large mainframe will become a file server for networked microcomputers. In addition, Toshiba has produced a laser disc that will store four gigabytes of information and they can gang four of these together on a microcomputer. That is a total of sixteen gigabytes of storage available through a microcomputer. Perhaps material will be moved from magnetic tape to laser disc. With the present advancements in optical scanners it may become feasible to put a lot of other library material onto laser disc.

All of this says to us that we must be ready to move with this tide or we will be swept aside by it.

Indiana University Libraries - Bloomington and Information Technology

"We are in an information age of dramatic growth and change. We want to create a shared vision of the future that builds upon our strong traditions of distinguished collections and superior services and that incorporates the technology-based information resources of the future....We envision a change in emphasis from the Libraries as a storehouse of information to the Libraries as a provider of information. We envision libraries that not only provide traditional access to collections but also deliver to desktop workstations information and documents in all formats through a network of campus and external resources." --Information Future*

The Libraries must take the leadership role at Indiana University in introducing new information technologies into the research process. The Libraries are not unique in their efforts to meet the challenges brought about by developing information technologies. The changes brought about by these challenges reflect those taking place in society as a whole. Technological advances in information and information access are a means to accomplish the mission of service to "One University" and are not ends unto themselves.

New information technologies are transforming library services. Library automation and electronic data already enable researchers to locate and use information faster and more effectively. As this technology evolves, researchers at their workstations will be able to access, acquire, and manipulate bibliographic, graphic, numeric, and full-text data.

The Libraries must continue their commitment to acquire information in traditional formats and maintain strong local collections. However, the Libraries recognize that they have never been able to acquire all research materials. Moreover, by adopting new technologies, the Libraries may choose in some cases to provide patrons with access to information rather than acquire it. Technology can help the Libraries to provide greater as well as more timely access to the information universe. Even though fiscal resources are inadequate, the Libraries must devote some share to obtaining information available in electronic formats. Regardless of format, the Libraries should provide information without user charges.

The goal of the Libraries is to maintain and improve the quality and scope of services by incorporating new technologies into the research process. In planning for implementation, the Libraries will consider the effect of these technologies on patrons and the successful use of the information resources in order to achieve this goal.

*Indiana University Libraries Strategic Plan

A Discussion Paper Prepared by Administrative Council. November 1988

To: Polley A. McClure
Dean, Academic Computing/University Computing Services

Carolyn A. Snyder
Acting Dean of University Libraries

From: Online Database Forum
Indiana University, Bloomington
Ann Bristow
Dave Lambert, co-chairpersons(1988/89)

Date: March 4, 1989

Attached is a report offering some background on the activities of the IUB Online Database Forum, some explanation of the issues as seen by the Forum and ending with several specific recommendations.

We would welcome discussion of this report in any way or by any groups you identify as appropriate.

cc: Forum participants

attachments:
Report of the Online Database Forum
List of participants in the Forum

March 3, 1989

Online Database Forum Report

Introduction

The Online Database Forum was formed in the fall of 1987 in response to a pressing need: to connect scholars and students with the online resources essential to their research and instruction. Participants in the Forum include technology specialists from the computing centers, information specialists from the libraries, and scholars. The Forum also has representatives from other committees dealing with important aspects of online resources, especially the Academic Computing Policy Committee (ACPC) and several of its subcommittees, one of which (Academic Information Systems) recommended the formation of a group resembling the Forum to address issues associated with determining what databases should be supported online. Our common purpose is to work towards facilitating the integration of appropriate academic information resources in an online academic information environment - an environment which we believe to be essential to support scholarship and research at Indiana University.

The Online Database Forum has been working within the context of other automation efforts at IU. The ACPC recently issued a recommendation that a large scale data server be available on the campus network to allow faculty and students to "access easily or extract and download smaller subsets of data for analysis on PCs, LANs, workstations, or departmental mini-computers." Since early 1987, a group of electronic library services called the Library Information and Reference Network (LIRN) has been available over the online academic information environment. These services include the ability to ask and receive answers to reference questions electronically, the ability to renew items or place them on reserve electronically, access to the OCLC cataloging utility database of 16 million records, and the ability to request electronically that books, article photocopies, and tables of contents be delivered to campus addresses. By January, 1990, the Libraries' card catalog will be available in computerized form and will be searchable from dedicated terminals at various locations on campus as well as through the network as a selection on the LIRN menu. The computerized catalog will provide electronic, distributed access to library holdings.

The Online Database Forum is concentrating its efforts on selecting new, or adding existing bibliographic, full text, numeric, or research databases to the network, and/or supporting the creation of databases locally. (Selecting could mean either leasing or buying tapes to load onto a local mainframe, or obtaining access electronically through a remote host mainframe).

The goal is to offer IU faculty and students an integrated electronic information environment through which they can have access to a broad range of computerized services and sources. In such an environment, a faculty member, from her or his office or home, will be able, for example, to search the Arts and Humanities Citation Index

database for citations on a particular subject or by a particular author, download the search results to a workstation and format them for inclusion in the bibliography of a research paper, send a copy of the paper to a colleague at another institution for comment, upload the search results to use as search criteria in the Libraries' online catalog, send the citations located in the online catalog to the campus delivery service and those not in the catalog to interlibrary loan, and query a reference librarian about alternative databases or search strategies. Or, a scholar could log onto the network to upload his or her data into an existing national or international database, or identify a data subset in a large database to download and analyze for incorporation into a report or article.

The Current Situation

A large amount of machine-readable data useful to scholars currently exists. Such machine-readable information, however, is of little value unless potential users are aware of its existence and, once aware, can obtain ready access. The Online Database Forum came into existence to address the problems of awareness and access by making a variety of databases accessible over a versatile network.

The scope of databases among which to choose is enormous. BRS, DIALOG and other vendors offer literally hundreds of bibliographic, full text, and numeric databases. Optical disk products containing bibliographic information, reference sources (e.g., the Oxford English Dictionary), technical manuals, numeric and statistical data, graphics, and the full texts of classic works are more and more plentiful. The U.S. Government publishes machine-readable census data and texts. Individual library catalogs (e.g., University of Illinois, Harvard, etc.) and shared holdings catalogs (e.g., OCLC and the Research Libraries Information Network or RLIN) are available in computerized form. Bionet, Geonet, and Humanist offer scholars online opportunities to communicate about and contribute to research in progress. Through Humanist, for example, humanities researchers have access to the Medieval and Early Modern Databank. This databank is a master set of 13,256 medieval currency exchange rate quotations compiled by Dr. Peter Spufford of the University of Cambridge for his Handbook on Medieval Exchange. It can be downloaded and used on any AT-class personal computer with ten megabytes of hard disk storage.

IU departments are involved in both the generation and use of bibliographic and nonbibliographic data. Some examples include: the Political Science Data Laboratory which provides access to the Inter-University Consortium for Political and Social Research based at the University of Michigan as well as the Indiana Poll; Business Computing Facilities and the Indiana Business Research Center in the School of Business which maintain financial, economic, and demographic data and offer access to INDIRS, the Economic Time Series, and the State Data Center for (Indiana) Census; the Swain Hall Library, which maintains a local preprint database; the Music Department which is involved in the creation of a database of scores and special collections; the Biology Department, where a biologist, Kathy Matthews,

obtained funding from NSF to develop and make available over Bitnet an international data bank on fruit flies; and the Chemistry Department, which includes the Molecular Structure Center that supports and services specific nonbibliographic data files.

At present, many of these databases must be accessed through a separate route, often with individual log on and search procedures; many are accessible to a limited, local group but not to a broader audience over the network; and, for many, data cannot be conveniently downloaded to a personal computer or workstation for manipulation. In addition, computing resources in terms of storage space and database server capabilities are insufficient to make versatile networking a reality. For example, the database management software and the disk storage space available locally are not adequate to support the computer science preprint database maintained by the Swain Hall Library or to allow the local database to interact effectively with comprehensive preprint libraries.

Why Does the Current Situation Need to Change?

Indiana University is beginning to fall dangerously behind in terms of computing, networking, and information technology compared to "early adopter" universities but also compared to peer Big Ten institutions. Carnegie Mellon's Andrew Project, which consists of a network, servers, and high-function workstations, is being used as the basis for the Mercury electronic library project. Mercury will demonstrate that a large scale, distributed electronic library can be built using existing technology, serve as a laboratory for studies on handling information electronically, and deliver to the desks of computer science faculty members a large percentage of the information they use. At Georgia Tech and at University of Southern California, the libraries and the computing centers have cooperated to make a variety of bibliographic databases available to faculty and students over the university computer network. CIC institutions, e.g., Ohio State, Michigan State, and University of Michigan, are in the process of networking bibliographic databases. At Ohio State, the ERIC educational database will soon be searchable in conjunction with the libraries' online catalog.

Two factors make falling behind particularly dangerous and point to the importance of local solutions to the problems of awareness and access: 1) computerized sources are searchable in ways that print sources can never be (for example, boolean searching, which allows concepts and terms to be combined, retrieves specifically defined sets of citations; searching for terms in full text databases allows relevant parts of texts to be rapidly and effectively identified); and 2) more and more computerized sources are available only in machine-readable form (for example, ARPANET provides a network for sharing research in progress that is unavailable in any other form; the U.S. Government is beginning to issue census data only in machine-readable form). Scholars, in other words, who do not have adequate awareness of and access to computerized sources will be hampered in their abilities to conduct research and will begin to have less of an impact on the

scholarly production of knowledge.

Recommendations

To ensure that Indiana University does not fall behind its peer institutions with regard to networking and information technology, the Online Database Forum has identified several issues that will need to be resolved before the long term goal of an integrated network of electronic services and sources can be achieved. Funding for database purchase, lease, access, and maintenance must be obtained through budgetary increases and reallocation: neither the IU Libraries nor the computing centers are adequately funded or staffed to do the work that needs to be done. Collecting more traditional library materials and providing computer clusters and classroom support are services the IU Libraries and computer centers cannot afford to sacrifice. The personnel to create and offer effective user support must be in place in both the libraries and computing centers or access to a wide range of services could lead to user frustration rather than satisfaction. Insofar as licensing, copyright, and charging requirements permit, databases housed at any of the IU campuses should be available on all campuses; cooperative funding mechanisms will need to be explored. A successful information environment for IU faculty and students will be an integrated environment. Data, records, and activities must be easily transferred between and among the services included in the Library Information and Reference Network (LIRN), the libraries online catalog (IO), and databases as they become available.

With these long term goals in mind, the Online Database Forum recommends that:

1) IU Libraries accept responsibility for funding the content portion of databases identified and selected through the selection mechanism already in place in the Libraries and/or suggested by the Online Database Forum (i.e., acquisition, flat fees, connect time, licensing, etc.), and IU Computing accept responsibility for funding the access portion of databases (i.e., communications, local front ends, network support, computing resources, etc.);

2) University departments be responsible for developing long range computing plans and for acquiring the hardware and providing the educational support necessary for faculty and staff to become computer literate;

3) During fiscal year 1989-90, IUB Libraries and Computing Services create and provide a procedure for maintaining and updating a searchable index/directory of databases and make it available to IU faculty, staff, and students over the AIE network. In the longer term, IU Libraries will explore the possibility of cataloging database information (e.g., databases owned or available anywhere within the University will be cataloged as are books and serials held by the Libraries) so that online sources can be located through searches in the Libraries' online

catalog;

4) During 1989-90 fiscal year, IUB Computing develop a large scale data server on the network that will be modular and extensible, support staging from tape to disk, support login for direct use of appropriate software, contain a variety of software (e.g., BRS search software, Ingres database software, SPIRES, etc.), support diverse networking protocols, and support optical disks;

5) During 1989-90, IUB Libraries and Computing Services acquire and mount a subset of the ERIC database on the AIE as a prototype to evaluate support demands, database software, and the capacity required to operate the large scale data server on the network. Such a prototype will serve to prepare the two service providers to acquire and mount databases of broader interest such as subsets of the Institute for Scientific Information's Citation Indexes during the following year;

6)Databases requiring per use charges be avoided whenever possible because of the unpredictable budgeting complications they produce. Recently, the Dow Jones Text database, which was offered over the network, was cancelled because its cost to IU was calculated on a per use basis. As usage increased and threatened to continue to increase, it became impossible to guarantee continued funding;

7) During the 1989-90 fiscal year, IUB Libraries and Computing Services undertake a pilot project to allow compact disk databases to be made available over the network;

8) The Online Database Forum continue to serve as the body that initiates new database projects and monitors the progress of projects already underway. Efforts should be made to attend to the interests of all database users and creators in the university community.

ONLINE DATABASE FORUM PARTICIPANTS

From September 1987, the following faculty members, librarians, and computing professionals have been invited to participate in meetings of the Forum; some have been regular attendees, some occasional, others have received information and participated as they could.

Miriam Bonham, Medical Sciences and Optometry Libraries
Ann Bristow, Library •@(1988/89)
Johann Dieken, BACS•
Dick Ellis, BACS •
David Farrell, Library •
David Fenske, Music Library •
Bob Goehlert, Library
Don Grinstead, UCS
Doug Grover, BACS •
Carol Hutchins, Swain Hall Library
Elton Jackson, Sociology Dept.
Dave Lambert, BACS •@(1987/88 and 1988/1989)
Scott Mantie, School of Business
Connie Miller, Library •
David Nordloh, English Dept. •
Rich Paustenbaugh, Business Library
Barry Rubin, SPEA
Carolyn Snyder, Library @(1987/88)
Sue Stager, BACS •
Brent Sweeny, BACS•
Barry Walsh, UCS
Alice Wickizer, Library •
Gary Wiggins, Chemistry Library •

•Principal participants in the attached Report and its recommendations
@ Co-chair

Collection Development Policy Statement:
COMPUTER-BASED INFORMATION SOURCES

I. Scope

The policy statement covers computer-based information sources acquired by the library through purchase, gift or subscription. It does not include those services for which users are charged a fee or products which are purchased solely for staff use (e.g. DIALOG for mediated search services or WordPerfect).

II. Definitions

Formats of computer-based information sources are defined as follows:

A. Machine Readable Data

Raw, unmanipulated numeric or textual information which must be processed by a computer. The data may be bibliographic, similar to a print product, whereby the information leads the user to another source, or non-bibliographic, in which the information, whether numbers, text, or images, is an end in itself. For the purposes of this document, machine readable data is categorized by its storage location.

1. On-Site Storage

a) Optical Disks/CD-ROM

These products generally reflect existing print reference sources. Hardware required includes a microcomputer, disk player, and printer. Existing standards for this format require a single workstation configuration which cannot be networked or accessed remotely.
EX: Public Affairs Information Service on CD-ROM

b) Microcomputer Diskettes

The information contained on a diskette is typically non-bibliographic, and it is usually coupled with a print counterpart, such as a monograph or journal issue. Hardware needed includes a microcomputer; applications software is needed to read the diskette. In general, a higher skill level is required for this category due to the lack of user friendliness and the use of applications software.
EX: MEMDB in prototype, Journal of Income and Wealth

c) Magnetic tape

The hardware required to read tapes is currently unavailable in the library. Bobst is not currently collecting tapes but social science data sets housed at

the Academic Computing Facility. are cataloged and entered in BobCat. Print codebooks are collected if the corresponding data set exists at the ACF.
EX: General Social Survey

2. Remote

a) Online Data

Files can be bibliographic or non-bibliographic and usually cover a specific subject area or discipline. The producer or publisher of individual files may be an association, government organization, or indexing service; vendors take the form of commercial intermediaries such as DIALOG, or non-profit entities, such as RLG or a university. Usage is typically interactive and accounted for by connect time. Hardware required includes a microcomputer and modem or communications gateway.
EX: DIALOG Business Connection; ARTFL

B. Software

In general, Bobst is not actively collecting in this area due to the lack of publicly-accessible hardware and the library's previously defined collection development policies. Categories in which collecting might take place include:

1. Instructional

This type of software guides the user through a series of concepts, processes, or models, usually in an interactive fashion. A typical application is automating a procedure or process in a specific discipline, such as teaching use of a Bunsen burner to students in an introductory class.
EX: Research Assistant, The Would-Be Gentleman

2. Applications

Applications software allows the manipulation, organization, or retrieval of textual or numeric input.
EX: Pro-Cite, SPSS

III. Selection Criteria

A. Content

The subject content and intellectual level of each information source will be reviewed in the context of the library's established collection development policy for that subject.

B. Selection issues - the following factors will be considered by selectors in consultation with the reference head

1. Product considerations

a) How does this source compare to existing print resources?

b) Does it contain additional information?

c) Could an existing print subscription be cancelled if this source were acquired?

d) Is the additional flexibility gained through automation significant for this particular source?

e) Does usage of print source merit an additional access point?

f) What is the coverage and scope of the information?

g) Is product capable of being accessed by simultaneous users?

h) Is product available through another library?

i) Is format in question the only one available for product? Is it the most cost effective/user friendly/accessible?

2. Vendor/contract considerations

a) Is a demonstration disk or trial account available?

b) Does the vendor have a reputation for producing high quality products with clear documentation?

c) Does the contract require restrictions such as obligation to return updated disks, guarantee of limited access, restrictions on downloading, definitions of users and/or uses of the information, liability from patron use of information, or restrictions on duplication of the documentation accompanying the product?

3. Hardware considerations

a) Is the hardware required currently available in the library?

b) How does purchasing the hardware from the vendor effect the pricing of the subscription?

c) If the vendor supplies customized hardware does it provide significant enhancement of the use of the product that could not be achieved with existing library hardware?

C. Implementation

The department head, in consultation with others as appropriate, will assess the impact of the proposed acquisition on:

1. Facilities

   Is additional hardware, wiring, furniture, etc. required?

2. Location of product/space

   Will renovation or reconfiguration of space be required?

3. Training

   What are the requirements for staff and users? What is the quality of the search software?

4. Service

   What is the impact on other reference services?

5. Technical support

   Is adequate expertise available in the library or on campus to support the use of the product?

col1254
5/5/89 md/jg

THE UNIVERSITY OF NORTH CAROLINA-CHAPEL HILL

DAVIS LIBRARY

The Machine Readable Data Files Center

Machine Readable Data Files, or MRDF's, are information or databases which are accessible through the use of a mocrocomputer. Davis Library's MRDF Center serves as a centralized location for this information, as well as the equipment and programs with which to access it. The MRDF Center, created in 1985, is based on a division of responsibility arranged by the provost in which the computation center has responsibility for Administrative Data Processing, the IRSS for mainframe data,the Microcomputer Support Center for general microcompuyter use, and the Library for data and information accessible through the use of a microcomputer. The current arrangement was recommended by a Library Task Force which studied the organization of this service. Administratively the MRDF Center is a part of BA/SS, but it collects microcomputer-based information resources in a number of different disciplines.

GOALS OF THE MRDF CENTER:

--To function as an information resource for the University community by providing information available in microcomputer-accessible formats.

--To help maintain Davis Library's role as a research library by collecting materials which, although outside the scope of traditional library print collections, are increasingly vital to scholarship. The MRDF may be seen as an experimental or prototype organization with the responsibility to stay abreast of the ongoing and often radical changes in the format and use of microcomputer-accessible information, which will have the capabilities needed to evaluate the technology and the various service issues involved in selecting and providing access to these rapidly-changing informational materials.

HOURS/ASSISTANCE: The MRDF Center is open Monday to Friday from 9-5 by appointment or (if not in use) by drop-in. Users needing assistance may contact Carson Holloway, MRDF Librarian (962-1151 or 962-8040). For those not needing special assistance, any BA/SS staff member is encouraged to open and service equipment.

CIRCULATION: MRDF's (disks or CD-ROMs) do not circulate or go out on ILL. In special cases exceptions will be made, but only through the MRDF Librarian. This policy is maintained partly in keeping with federal copyright laws, with software licensing laws, and with University policies

which forbid the unauthorized copying or use of computer software, including in some cases the use of unauthorized equipment to access software. The Library has a legal and ethical responsibility to control as far as possible the circumstances under which materials could be copied or unlawfully used. A second reason for the non-circulation policy is concern for data integrity; loss of control over circulated materials could lead to data being wilfully or accidentally changed or erased.

ARRANGEMENT OF MATERIALS: The materials collected in the MRDF Center include the following:

--Software not catalogued and available for use with databases (such as Lotus 1-2-3).

--Catalogued materials: 1) Print materials (such as manuals) without an accompanying MRDF are given an LC classification number, 2) Print materials with accompanying MRDF(s) are catalogued under the accession system, and 3) MRDF's with no accompanying print documentation are also catalogued under the accession system. The accession system label conveys the following information: the collection (in this case MRDF), the format (such as CD-ROM), whether the material is a monograph or a serial, and an accession number.

SELECTION OF MATERIALS: Suggestions for purchase are made by the MRDF Librarian. Suggestions are made according to the following principles: 1) the possible/probable level of use of the database, 2) if the database is an exploration of possible technology, a groundbreaking innovation, or a likely prototype of future developments, 3) if the database, through the exploitation of technology, improves on non-electronic resources by providing more information, better and faster access to information, the capability of organizing and packaging information, etc., and 4) if the database is highly recommended by knowledgable sources (colleagues, reviews or articles, faculty.) Materials are purchases on special funds rather than being a part of the regular book budget.

EQUIPMENT:

--1 IBM AT with a math coprocessor.
--1 IBM PC with a 20-megabyte hard disk.
--1 IBM PC with two 360-K floppy disks.
--1 Hewlitt-Packard LaserJet Series II printer.
--1 Epson FX 280 dot-matrix printer.
--1 IBM graphics printer.

STATISTICS: Administrative statistics on the use of the MRDF Center are maintained on a continuing basis. Data gathered include the date of use, the academic status of the user, University department affiliation, databases and program files used, and amount of elapsed time in the MRDF Center. Sheets to be filled out for MRDF users are provided

in the second drawer of the MRDF file cabinet. Sheets are completed by library staff and placed in a folder in the second drawer of the file cabinet. CD-ROM statistics for U-Search are kept by a computer program; users answer a series of questions before they can access the database. Data are compiled on a quarterly basis for the BA/SS annual report.

USE OF THE LASER PRINTER: BA/SS staff (and other library staff) may use the laser printer on a sign-up basis (a calendar for sign-up can be found in this notebook.) The IBM-AT is not available for general word-processing; it must only be used by staff to print. The order of priority for use is as follows: 1) patrons, 2) MRDF Librarian and assistant, 3) library staff. Please note that patron use will take precedence over staff use even in cases of prior sign-up.

UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL

Collection Development Policy for Machine Readable Data Files On Disks

June 26, 1985
revised July 1, July 3, September 24, 1985

Introduction:

The fact that access to computer technology is now widespread within society as a whole has led to a corresponding increase in information resources which are available in new formats such as microcomputer floppy disks (also referred to as "diskettes") and magnetic tapes. While these new formats differ significantly from those conventionally collected by libraries, the information that is carried or conveyed through this new medium, whether bibliographic, numeric or textual, constitutes a logical extension of those materials which libraries currently collect. As one example of the types of information now available in the new format, in the past the Chronicle of Higher Education developed various yearly surveys and statistical analyses which were published regularly in hardcopy in issues of the Chronicle. These tables and surveys are now available for purchase on floppy disks, which allows the individual user to manipulate and reformat the data in whatever manner and for whatever purpose the user may choose. Other types of information, especially statistical information, census data, etc. which previously appeared in book or serial format, have likewise become available on disks or magnetic tape. The U.S. Bureau of the Census publications, The County and City Databook, County Business Patterns, and Population Estimates are examples of titles formerly in hardcopy which are also now available as data diskettes. Examples of non-numeric data files are: CareerScan IV, Directory of Public High Technology Corporations and the Academic American Encyclopedia. Examples of instructional data files are: Ethical Problem-Solving in the Social Sciences and Harmonic Analysis and Dictation. Increasingly, other information of an instructional, reference and research nature are also becoming available on magnetic and optical disks and in other machine readable formats.

While policies on the selection and acquisition of books, serials, videotapes, audio tapes, films, and filmstrips currently exist, few policies have been developed to adequately handle the selection and acquisition of data diskettes and datatapes. The following document is intended to provide such a policy for magnetic and optical disks only. Both fixed and manipulatable types of data files in these categories are covered by this policy. Because of the special requirements which apply to magnetic tapes, including their often being accessible only through mainframe computers, magnetic tapes are excluded from this policy. It should be noted that, due to the rapid and unpredictable technological advances which are taking place in terms of computer applications, this policy for data diskettes will necessarily require frequent review and revision.

1. General Considerations

A. The **BASIC PREMISE** in establishing a formal policy on the selection and acquisition of materials in a machine readable format is that an academic, research library has as its primary responsibility the provision of access to the most extensive and most sophisticated sources of information available within the limits placed on the library by its financial, personnel and physical resources. Materials, such as data diskettes, which provide access to and enhance the use of significant information resources are appropriate additions to the library's selection and acquisition programs. They add a new dimension to the information which a library can provide and therefore, subject to financial and physical contraints, belong in the library's collections.

B. Because of the existence of a multitude of materials which appear in machine readable formats some means of differentiating among these formats for the purpose of selection and acquisition is necessary. Any **DEFINITION** of materials to be covered by this policy must take into account the existence of two basic types of disks (diskettes): 1) Program diskettes; and, 2) Data diskettes. Program diskettes contain "a logically arranged set of programming statements (or instructions) defining the operations to be performed by a computer in order to achieve the desired results." Examples of program diskettes are: Lotus 1,2,3, dBase III, Multimate, VisiCalc, SuperCalc, R-Base, etc. Data diskettes contain "characters grouped together in specific patterns, to which meaning is assigned: commonly used to designate the numbers, facts, concepts, or the like to be processed by a program." Examples of data diskettes are to be seen in the aforementioned U.S. Bureau of the Census publications, The County and City Databook, County Business Patterns, Population Estimates, CareerScan IV, PsycLit, Ethical Problem-Solving, etc.

C. Included within the **SCOPE** of this policy are only issues related to the selection and acquisition of data disks. Issues dealing with the selection and acquisition of individual program disks are excluded from consideration in the present policy, except in those instances where the individual publisher or supplier has included them as an integral and inseparable part of the data disk. (At present the selection and acquisition of program disks is expected to take place outside of the library materials budget. The fact that program disks are presently excluded from this policy is not meant to preclude their possible inclusion, contingent upon circumstances, at some later date.) Further, data disks must be located in those areas within the Academic Affairs Library which have the ability to handle and service these materials. **DATA DISKS OF ANY SORT CANNOT BE HOUSED IN OPEN STACKS IN ANY ACADEMIC AFFAIRS LIBRARY UNIT.**

2. Selection Procedures

The same selection CRITERIA which are currently applied to the selection of books, serials and standing orders should be applied to the selection of data disks: 1) The materials enhance the library's collections; 2) The material is of a high quality both physically and intellectually; 3) The materials are within the subject parameters of the collections; and, 4) The materials are appropriate to the teaching and research needs of the community which the library serves; and, 5) In light of items 1-4 above, the cost of the material must be weighed against the cost of other possible purchases. Finally, in the case of selection by academic departments, it must be emphasized that data disks selected must be cataloged and housed within a unit or Department of the Academic Affairs Library. (They cannot under any circumstances be indiscriminately housed in the library's main stacks.)

As is the case with books, serials and standing orders, academic departments cannot use library materials funding to select materials to be held outside of the library system. In addition to these preexisting criteria for selection, several other factors must be considered in the selection process:

A. The data disks must be available for acquisition by the library. Agencies which sell or provide only to individuals rather than to institutions should be excluded from the selection process. Under no conditions will blank disks be provided in barter by the library as a basis for any agency providing or producing data disks.

NOTE: If an individual who has acquired or been given materials wishes in turn to donate these materials to the library with the understanding that these materials will be available generally, the materials should then be treated as would any other unsolicited gift.

B. All hardware, i.e. "the physical equipment and components in a computer system," required to operate the data disk must already be available either in the Department selecting the material or in the area which will ultimately house the data disk.

NOTE: If materials are being selected by a Departmental Library and are to be housed and used in that Departmental Library, then the Departmental Library must already possess the hardware necessary to run the data disks. Materials selected by an external unit for possible housing in a Departmental Library or any Academic Affairs Library unit must first gain permission and approval of that unit before selection. The library or unit concerned will determine whether it is interested in having the material, and then if it has the capability for housing and servicing the material requested.

C. All programs required to operate the data disk -- except those which are an integral part of the material under consideration for selection -- must already be available either in the Department selecting the material or in the area which will ultimately house the data disk. Program disks which are not available to the public, e.g., those used for internal departmental record-keeping or other non-public operations, do not meet this requirement.

NOTE: If materials are being selected by a Departmental Library and are to be housed and used in that Departmental Library, then the Departmental Library must already possess the programs necessary to run the data disks. Materials selected by an external unit for possible housing in a Departmental Library or any Academic Affairs Library Unit must first gain permission and approval of that unit before selection. The library or unit involved will determine whether it has the program capability of housing and servicing the selection.

D. Part of the selection process must also include the prior determination, i.e., before ordering, that sufficient personnel with an adequate level of training in hardware and software use are available in order to assure access to and proper use of the materials selected.

NOTE: If materials are being selected by a Departmental Library and are to be housed and used in that Departmental Library, then the Departmental Library must already possess the personnel necessary to run the data disks. Materials selected by an external unit for possible housing in a Departmental Library or any Academic Affairs Library unit must first gain permission and approval of that unit before selection. The library or unit involved will determine whether it has the personnel capable of servicing the materials selected.

3. Funding

Funding for those materials which meet all the selection criteria above will be the same as currently exists for books and new subscriptions. One time purchases of data disks will be considered as being the same as an expenditure on books. Subscriptions to services which provide data disks on a recurring basis will be considered as being the same as a New Subscription. There will be no special allocation for the purchase of these materials, and they must fall within the financial limits provided by the existing budget allocations. The same limitations which currently exist for Book and New Subscription expenditures will also apply:

A. Individual data disks being requested for purchase within a Book Allocation and costing over $250 but less than $500 must be approved by the University Bibliographer before they are forwarded to the Acquisitions Department.

B. Individual data disks being requested for purchase within a Book Allocation line and costing more than $500 must be approved by the Research Fund Committee of the Library Administrative Board before they are forwarded to the Acquisitions Department.

C. More expensive data disks can be considered by the Research Fund Committee at both its fall and spring meetings. These requests must meet the same criteria and standards which apply to all materials to be considered by the Research Fund Committee.

D. Requests for additional funding in the Book and New Subscriptions lines for the specific purchase of data disks should be included with the budget requests which are submitted in May/June of each year. Although the amount should be included as part of the Book or New Subscription allocation request, as is appropriate for the type of material, the amount requested for data disks should be specifically and separately stated.

E. Expenditures for data disks will be monitored by the Collection Development Department on a regular basis in order to determine the increase or decline in interest in the acquisition of these types of materials; their impact on existing collections and services, will also be studied.

4. Acquisitions

As is currently the case with Books and New Subscriptions, the appropriate order card should be generated by the unit ordering the material. Information on the order card should include the specific title of the data disk and all information appropriate to identify the agency selling the material. It should be clearly indicated on the order card that the material is a data disk. The Academic Affairs Library location for the data disk must be indicated. If the location for the material is other than that of the individual approving the order, then the order card must also contain the signature of the individual who is in charge of that location. No materials will be accepted or processed for acquisition unless there is a clear indication that the location of the material has been approved.

5. Cataloging, Circulation, Copying and Use of Data Disks

Policies concerning cataloging, availability for use by the public, circulation and copying in addition to policies regarding public access to all the programs, hardware and data disks covered directly or indirectly by this policy are presently being developed. Upon approval these policies will be appended to this document.

6. Special Considerations

It is a stated and basic principle within this policy that no data disks should be stored in the open stacks of any Academic Affairs Library unit. Also included within this category of exclusions are any floppy disks (diskettes), magnetic tapes or other machine readable materials contained physically within books or serials. The main stacks provide neither access nor security for any of these materials. Those disks, especially ones contained in books, which are inadvertently checked out through the circulation system suffer the immediate and complete physical loss of the information stored on them. Dust, dirt, heat, humidity, compression, and improper storage and handling will also result in disk failure. A controlled environment is absolutely necessary in order to provide even minimal disk preservation. A location other than the main stacks must be established and approved before these materials are cataloged. Any materials which are not accepted at locations other than the main stacks must be either returned to the vendor or publisher, or discarded.

Because of the difficulty which the Library of Congress has had, until recently, in establishing cataloging procedures for data disks, tapes, program tapes, etc., especially those which are enclosed in books, we have acquired a number of unexpected data disks and programs which were included in books. These should be returned to the individual who ordered the material so that he or she may establish their specific library location in light of the above policy. In the future any such material identifed after receipt should be shown to the individual who ordered the material so that proper disposition of the material can be arranged in accord with this policy.

As noted previously in the policy, blank disks will not be provided nor can be acquired through the use of Library materials funding for exchange with an agency or publisher for data disks.

The Ohio State University
Collections Advisory Council
DRAFT
April 18, 1988

CD-ROM Technology Policy Recommendations

As an academic research and teaching institution, the Ohio State University has an obligation to expose its students and faculty to the cutting edge of technology. The University Libraries have always led the way in the application of technology to information retrieval. Although some critics argue that CD-ROM may be a passing technology, it will undoubtedly be replaced by another more sophisticated technology. OSUL cannot afford to wait for the technology to be standardized before proceeding. In fact, as Attachment A reveals location libraries in conjunction with departmental faculty have already introduced a number of CD-ROM databases on campus. However, the Libraries have not as yet been able to undertake a systematic introduction of this technology as a result of limited funding.

Training

In some respects the issue of training for CD-ROM technology has not been adequately addressed. Individual location libraries have instituted user education programs geared toward the individual databases they provide. Since the number of databases available on campus has increased, centralized training in the concept of Boolean searching rather than the specific tools may be justified. Although the systems are user-friendly, the potential for end-user searching and locally mounted databases as well as CD-ROM technology compels the libraries to consider centralized instruction.

Criteria

CD-ROM databases should be evaluated against the following criteria when purchase decisions are made.

1) Is the subject focus of the database relevant to the subject collection of the location requesting the title? Is this the most appropriate location for the first copy of the title?
2) Does the technology provide improved access to the information, i.e., Boolean searching, keyword searching?
3) Is the software used to access the database of high quality? Does it offer the accepted minimum criteria such as field qualification and adjacency? Can users be expected to use the system without extensive training?
5) The cost of the product should be evaluated in terms of the cost for other titles and the number of potential users. Is the anticipated use more cost effective than subsidized searching? Is the use so significant as to warrant locally mounting the database? Trial periods are suggested whenever possible and location libraries should be encouraged to keep use statistics.
6) Does the database provide improved service to patrons in addition to saving staff costs? Does having the CD-ROM available make ready reference and other activities more efficient? For example, would the CD-ROM version of Books in Print for technical services reduce processing costs?
7) The purchase of unique databases vs. duplicate titles must be considered and justified in the evaluation process.

Cataloging

The Workgroup recognizes that one of the underlying issues in setting the cataloging policy for CD-ROM technology is the method of acquisitions. Until adequate funding is available to purchase all titles on subscription, minimal level cataloging should be examined as a potential solution. The principle need at this time is to provide access for public service locations; minimal level cataloging would provide this access. In addition, Attachment D from Ichiko Morita to Carol Hawks addresses the issue of revised software versions. The latest version note seems to be the best solution. Departmental Libraries are encouraged to use LCS online maintenance to add the last software received if a subscription is cancelled.

Equipment Compatibility

Equipment purchased for the use of a CD-ROM product should be, whenever possible, that which is specified by the distributor of the data base. In most cases the distributor will specify a microcomputer and CD-ROM drive to access the data. (It is highly unlikely that a CD-ROM distributor would choose a hardware configuration that is outside the "mainstream" of the personal computer industry.) In the event that the distributor does not recommend specific computer equipment, the librarian directly involved in the purchase should specify the hardware based on available information and location needs.

Retention of Paper Copies

There are instances where paper copies of a CD-ROM do not exist, e.g., ABI/INFORM. However, in most cases a paper version of a CD-ROM is available. As a general rule, the paper version should not be cancelled. There are several justifications for this rule: one workstation cannot replace the printed version, some CD-ROM services are leased and the database must be returned if cancelled, the long term future of CD-ROM technology is unknown. In general, the cancellation of a paper version should be considered as a separate issue regardless of the existence of a CD-ROM version. Such cancellation should be reviewed by the Collections Advisory Council.

THE STATE UNIVERSITY OF NEW JERSEY
RUTGERS

Center for Computer and Information Services • Hill Center for the Mathematical Sciences
Busch Campus • P.O. Box 879 • Piscataway • New Jersey 08855-0879

RECEIVED
DEC 14 1988
ADMINISTRATIVE OFFICE
R. U. L. LIBRARIES

December 8, 1988

TO: Gary Golden
Camden Library

FROM: Dr. Leny Struminger LS

SUBJECT: Checkout Policy for Computer Software

Following our phone conversation and due to several requests received from faculty at Newark and Camden, I am writing to let you know about a temporary lending policy for our software. We decided on one to one basis to allow programs to be checked out to faculty for a short period of time for evaluation purposes only. In order to obtain the program and accompanying documentation a form which is enclosed, must be submitted to us by faculty indicating what package they want to test and to agree that none of the material will be copied for future use.

This is a temporary policy to see what interest exists for the Center software and to be able to open access to all Rutgers faculty to our collection. At this time we own only one copy of each program and this is the only way we will be able to circulate the material.

Thanks for your cooperation on this matter. Please do not hesitate to call for additional information. I can be reached at extension 2427.

Attachment

cc E. Michna
Peter Graham
jc

SOFTWARE INFORMATION CENTER

FACULTY BORROWER FORM

This software package, _____, package # _____, is the property of CCIS Software Information Center at Rutgers University. This package and other collection in the Center are made available to faculty under the following conditions:

1. The software is to be used only for the purpose of demonstration or evaluation (possibly for acquisition).

2. The software and accompanying materials may not be copied, decompiled, modified in any way (except in accordance with installation instructions in the documentation), or loaned to third persons (except for teaching assistants or faculty involved in teaching the same courses as the borrower). The software may be copied onto hard disk if that is necessary to execute the program; however, any such copy must be erased at the end of the demonstration or evaluation session.

3. The software package, including accompanying materials, must be returned completely intact to the Software Information Center **no later than two weeks** from this date.

Since Rutgers has obtained the use of this software package by express agreement with the publisher, it is a university requirement that all faculty and other employees adhere strictly to above terms.

Today's Date: _____

Name: _____ Signature: _____

Dept: _____

Telephone: _____ 59

spec
OMS
Systems and Procedures Exchange Center

CATALOGING POLICIES AND GUIDELINES

Cataloging MRDF at UF
A Revised Policy Proposal

Prepared by Suzy Shaw

1. **PURPOSE:** In order to improve access to the machine readable data file (MRDF) collection maintained in the Tape Library, cataloging records for selected data sets will be created for inclusion into the FOCUS online catalog and the national bibliographic networks.

2. **SCOPE:** Only data sets owned by the Library will be cataloged.No MRDF lacking documentation will be considered. Data sets on tape, cartridge, diskette, and CD-ROM will be included.

3. **PRIORITIES:** Cataloging priorities will be set by Reference staff who deal with MRDF's regularly.

4. **STANDARDS:**

   A. Cataloging will be performed according to existing rules and formats with the exceptions listed below. At present, this means the OCLC MRDF format, 2nd ed., the NOTIS MRDF format and the draft edition of Revised Chapter 9, AACR2.

   B. Description will be based on documentation. With the exception of microcomputer software where hardware and system requirements are readily available, no magnetic media will be mounted for cataloging purposes.

   C. LC Subject headings and classification numbers will be assigned.

   D. Access to Reference and Systems numbering schemes will be provided using the stock number field.

   E. The presence of backup copies will be assumed. Backups will not be recorded on copy holdings records

   F. The MRDF format accommodates both monographic and serial treatment. Until detailed guidelines are developed, this will be up to the judgement of the cataloger.

   G. Volume holdings information will be provided where appropriate.

   H. The details concerning problems of seriality, "loose-leaf" analogy, ownership, and record maintenance have yet to be resolved.

5. **FIELD USE POLICY:**

A. Chronological information will be coded in 045 for future index capability. If required for clarity, it will also be given in free text form in note field 523.

B. Note fields for type of file or data (516) and file characteristics (565) will not be used. This information will be recorded in the form of a 256 field.

C. Because data tapes are owned by and stored in the library, physical description in the form of a 300 field will be included. Unless the medium is specifically stated to be CD-ROM or floppy disk, computer tape reels will be the presumed medium.

D. Notes about other formats available will not be included.

E. Linking field complexity notes will be included when an important and non-obvious relationship with another work exists.

F. Detailed summary notes will be constructed by Reference Staff, OR, suitable quotations from documentation will be indicated by them for inclusion in the 520 field. Display constant "Summary" will be used for locally constructed notes only.

G. Contents notes will not be included unless essential to clarify holdings.

6. INPUT:

A. A preliminary record will be input on FOCUS. MRDF cataloging workforms will be input into OCLC. Records will enter the FOCUS online catalog by means of the ONI bmer interface. Records will later become part of the RLIN database in the course of routine tape loading.

B. Although full level records have been created, they will be entered into OCLC as Level K to permit ease of upgrading, especially of the 256 information. (This course of action is recommended by OCLC.)

C. MRDF records will reside in the regular LTUF processing unit of FOCUS and not the LTTL processing unit. Because there is no location defined for the tape library, a universal message on the copy level note area will state: Shelved in Tape Library; consult Reference staff.

62

mm3 4-25-88

MRDF Cataloging Worksheet

Part I

cus #_____ Brief title_____

'stems doc # 035 (FU)SYS _____

stems TL # 037 Sys _____

XED FIELD

.C: NEW Rec stat: n Entrd 880000 Used: 880000
ɔe: m Bibl lvl: _ Govt pub: Lang: N/A Source: d Frequen: _
.e: u Enc lvl: K Machine: a Ctry: ___ Dat tp _ Regulr:
ɜc: a Mod rec: Dates: ____ ,

FUG $c FUG 049 FUGG

---

entification numbers

erence Shelflist: 037 Ref _____

ck number: 037 _____

er numbers:

ɔnological Coverage

045_____

---

es and statements of responsibility

--

01

01

01

---

int information

63

Collation 300

---

500 Description based on:

521 8 Audience:

506 Restrictions on access:

Summary:

538 System requirements: IBM 370 or equivalent mainframe.

other 5uu's

Subject Headings

6__ __

6__ __

6__ __

---

call number

Added Entries

7__ __

7__ __

7__ __

Cited in:

File Characteristics

256

Variables

565 [no. of variables] $a_____

$b_____

$b_____

$b_____

$b_____

[unit of analysis] $c_____

[data universe] $d_____

[filin scheme/code] $e_____

Links to other materials

65 FOCUS #_____

THE STATE UNIVERSITY OF NEW JERSEY
RUTGERS

Interdepartmental Communication

11/9/88

RECEIVED

NOV 14 1988

ADMINISTRATIVE OFFICE
R.U.L. LIBRARIES

To: Peter Graham
AUL/TAS

From: Samson Soong
Head, BCD

Re: Software cataloging project

The software cataloging project has been in operation since 9/1/88 and continues to run smoothly and efficiently.

In addition to 81 Wisc-Ware software, a total of 107 Kinko's items have been received by the Software Information Center. As of this writing, all of them have been cataloged and entered into Geac. A copy of two Kinko's records are attached for your information.

cc: J. Boyle
M. Gaunt
B. Ginder
E. Greenberg
L. Struminger (CCIS)

66

04 Geac Marc Data Entry System Date: 88-11-11
Function: DISPLAY

SEARCH TYPE: a KEY: kinko's

REF#		RSN	KEY	TITLE
1	m	00839856	Kinko's Service Corporation.	DEGraph
2	m	00839857	Kinko's Service Corporation.	Differential equations
3	m	00838694	Kinko's Service Corporation.	Drexel plot
4	m	00838301	Kinko's Service Corporation.	Drill
5	m	00840163	Kinko's Service Corporation.	Electromagnetism
6	m	00838315	Kinko's Service Corporation.	Energetics and metabolis
7	m	00838294	Kinko's Service Corporation.	Escher-Sketch
8	m	00838348	Kinko's Service Corporation.	Event tutor
9	m	00838319	Kinko's Service Corporation.	Evolution
10	m	00838625	Kinko's Service Corporation.	FEMG: a finite element m
11	m	00838350	Kinko's Service Corporation.	ForthTalk kernel
12	m	00838351	Kinko's Service Corporation.	ForthTalk library
13	m	00839858	Kinko's Service Corporation.	FunPlot
14	m	00839953	Kinko's Service Corporation.	FunPlot - 3D
15	m	00838323	Kinko's Service Corporation.	General chemistry multip
16	m	00839824	Kinko's Service Corporation.	GeoStructures

Select Number: n DIR: + AMT: 01

SEARCH TYPE: a KEY: wisc-ware

REF#		RSN	KEY	TITLE
1	m	00840424	Wisc-Ware.	ANTS - Reliability theor
2	m	00833162	Wisc-Ware.	API
3	m	00833124	Wisc-Ware.	Applied basic physics -
4	m	00840409	Wisc-Ware.	ASA - Analysis and simul
5	m	00833163	Wisc-Ware.	ASK
6	m	00832432	Wisc-Ware.	CAFE - Connection arrays
7	m	00841352	Wisc-Ware.	CIRC - Electronic circui
8	m	00833134	Wisc-Ware.	Controlling the IBM 6 pe
9	m	00832376	Wisc-Ware.	CRITPATH
10	m	00841354	Wisc-Ware.	CRYSTALLAB - X-Ray diffr
11	m	00831040	Wisc-Ware.	Data transformation
12	m	00841341	Wisc-Ware.	Differential equations a
13	m	00832858	Wisc-Ware.	DTREE
14	m	00832374	Wisc-Ware.	Duke language toolkit
15	m	00832637	Wisc-Ware.	EDC - Elementary digital
16	m	00840427	Wisc-Ware.	EPIGRAPH - Plant disease

Select Number: n DIR: + AMT: 01

67

```
GRSN : 00840167   DATE:  88-11-C1

Original Entry Date : 04  88-11-C1    Updated date          : 00  CC-00-00
Print on proof list :  d  88-11-C1    Completed for load    :  y  88-11-01
Transfer to Circ    :  y  88-11-C1
New Additions list  :  y    Marc request          :  n    Aged record  :  n
Record on order     :  n    Generate Cards        :  n    Print Labels :  n
Remove in Circ      :  n    Non-Circulating       :  n

  100   1       1 $aRuppeiner, George.
  245   1       10$aOptics lab$h[computer file]
  250   1         $aVersion 1.C.
  260   1         $aVentura, Ca. :$bKinko's Service Corp.,$c1988.
  300   1         $a1 computer disk ;$c3 1/2 in. +$e1 manual.
  500   1         $a"Physics Package No. 090."
  538   1         $aSystem requirements: Macintosh computer, minimum 512K, with Finder, version 5.3.
  690   1        4$aPhysics.
  710   1       20$aKinko's Service Corporation.
  753   1         $aMacintosh.
  966   1         $c1$1SOFTW$aKinko's090$aSTACKS$b39030013594900
```

```
GRSN : 00840158   DATE:  88-11-C1

Original Entry Date : 04  88-11-C1    Updated date          : 00  00-00-00
Print on proof list :  d  88-11-C1    Completed for load    :  y  88-11-01
Transfer to Circ    :  y  88-11-C1
New Additions list  :  y    Marc request          :  n    Aged record  :  n
Record on order     :  n    Generate Cards        :  n    Print Labels :  n
Remove in Circ      :  n    Non-Circulating       :  n

  100   1       1 $aKramer, Richard.
  245   1       10$aHeavenly Mac$h[computer file]
  250   1       . $aVersion 1.C.
  260   1         $aVentura, Ca. :$bKinko's Service Corp.,$c1988.
  300   1         $a1 computer disk ;$c3 1/2 in. +$e1 manual.
  500   1         $a"Physics Package No. 086."
  538   1         $aSystem requirements: Macintosh computer, minimum Macintosh Plus. Hypercard is also required
  690   1        4$aAstronomy.
  710   1       20$aKinko's Service Corporation.
  753   1         $aMacintosh.
  966   1         $c1$1SOFTW$aKinko's086$aSTACKS$b39030013594868
```

The State University of New Jersey

RUTGERS

The University Libraries

Technical and Automated Services August 11, 19

To: Gene Young, Associate Vice President for Computing Services

From: Peter Graham, Associate University Librarian for Technical and Automated Services

Subject: Access to the Electronic Software Library

I'm enclosing a copy of the guidelines for cataloging material for the ESL; it amounts to our cataloging manual at the moment. It is the product of useful work primarily by Sam Soong, the Head of our Bibliographic Control Department which will actually do the work (he has really thrown energy into this project); and of a good deal of useful consultation within the Libraries and from Kathy Ciociola and Gert Lewis.

I'm about to go off on vacation for a couple of weeks but I wanted to get this onto your desk for your return. We expect to have an actual person in place to do this fairly shortly, and we're aiming at September 15. When we meet on the 29th of August we'll all know more. We'll be working from the printed catalog materials; we'll all hope for actual arrival of the materials so the access can be real.

See you soon. I'm delighted we're able to work so well together on this.

cc: Peggy Seiden (Carnegie Mellon)
Steve Gilbert (EDUCOM)
Dick Storer, Head, OCIS

Joanne R Euster, University Librarian
Samson Soong, Head, BCD

THE STATE UNIVERSITY OF NEW JERSEY
RUTGERS

Interdepartmental Communication

August 10, 1988

CC: Jeanne Boyle
Kathy Ciociola
Marianne Gaunt
Bernice Ginder
Evelyn Greenberg
Gertrude Lewis

TO: Peter Graham
Associate University Librarian for Technical and Automated Services

FROM: Samson Soong
Head, Bibliographic Control Department

RE: Cataloging of the Software Library

Following are agreements re cataloging of items at the Software Library that we reached at the meeting of 8/9/88. Present at the meeting were J. Boyle, K. Ciociola, M. Gaunt, E. Greenberg, P. Graham, G. Lewis, and S. Soong.

1. An action plan utilizing brief record cataloging to provide access to the material at the Software Library by September 15, 1988 is considered as an interim measure, until a decision is made with regard to the level of cataloging needed on a permanent basis for machine-readable data files on campus (e.g. items not owned by the libraries).

2. Brief record will be created on GEAC in its MRMS sub-system. Training and work be initially limited around GEAC system (i.e. no RLIN search for now).

3. A LOC/agency code "SOFTW", for Software Library, will be established. The existing material type "STACKS" will be used. New material types may be established when there are sub-collections at the Software Library.

4. The software and any accompanying user manual/guide cataloged will be assigned one barcode number. We will not actually barcode the item.

5. Accession numbers will be used for items to be cataloged. The following sequences will be used for Kinko's, Wisc-Ware and CONDUIT software respectively: K--, W--, C--.

6. For brief record cataloging of these items, the following data elements will be included:001,100,245,250,260,300,538, 690,700,753 and 966. (see the appendix for detailed information for each field and a brief record example.)

7. The prescribed data elements will be prepared by using primarily the publisher's catalog and accompanying documentation,e.g. user's manual. (The chief source of information rule for computer files, i.e. the title screen(s),in AACR II is to be ignored for creating these temporary records)

8. Authority work will not be done on these tentative records.

9. Bibliographic maintenance of these records, including holdings information, will be kept as minimal as possible. We will appreciate hearing of withdrawals.

10. A graduate assistant will be hired and trained in BCD to create brief records on GEAC.

11. Software and other computer files owned by the libraries will continue to be cataloged via RLIN. In the event of duplication, holdings (i.e. 966s) for these library-owned items, for the time being, will not be attached to the brief records created as part of this project.

12. All concerned people, including potential major users of the software in the university community, will be notified of this project.

13. Brief record cataloging will be done on a trial basis and be evaluated at the end of 12 months.

APPENDIX
Data Elements to be Included In the Brief Record

(The rules for the descriptive cataloging of machine-readable data files published in the revised chapter 9 of the AACRII will partially be followed. It is impossible to list down at this point all the exceptions to the rules as they will evolve once the project begins.)

a. 001 — enter CSN in the following format: XXXXnnnnnnnn-D, where the n's are the last eight digits of the barcode used. The prefix "XXXX" will be determined and will be for items at the Software Library. (Note: "D" is used by RLIN for computer or data files)

b. 1XX — enter the author's name, as found in the publisher's catalog or the accompanying documentation, regardless whether it is a personal author or a corporate body (Avoid choice of main entry decision). In the case of a personal author, enter the author's last name, a comma and a first name, and if present, a middle initial. In the case of more than one author, enter the first one here and the others as 7XXs.

c. 245 — Title (Includes only the title proper. Excludes parallel title, other title information and general material designation. Excludes also statements of responsibility , e.g. designed by Donna Stanger, programmed by Lon Koenig)

d. 250 — Edition statement (transcribe a statement relating to a named reissue of a software, or to an edition of the software that contains differences from other editions, e.g. Version 1.2. Excludes statements of responsibility relating to the edition, e.g. program has been converted from BASIC to FORTRAN by Allen P. Smith)

e. 260 — Publication/distribution, etc., area (Includes place of publication, distribution, etc.; name of publisher, distributor, etc. and date of publication,, distribution, etc. Excludes statement of function of publisher, distributor, etc.)

f. 300 — Physical description area (Includes extent of item, e.g. 2 computer cassettes, but excludes other physical details, e.g. single sided, single density, soft sectored; Includes also dimensions, e.g. 5 1/4 in., and accompanying material. Excludes details of accompanying material, e.g. 52 p. : ill. ; 20 cm.)

g. 538 — System requirements (always includes a note on the system requirements of the software if the information is readily available. Precede the note by "System requirements").

h. 690 — Subject heading (One subject heading will be assigned. In establishing subject for the item, follow the form in the list of headings that has been agreed on).

i. 7XX — added entries (When there are more than one person or corporate body sharing the responsibility for the software, enter the other names here)

— Make also an added entry for one of the following, as appropriate:

Kinko's Service Corporation
Wisc-Ware
CONDUIT

j. 753 — Technical details access to MRDF (Includes make and model of machine and operating system if the information is readily available or implicit; excludes programming language).

k. 966 — Copy information (one 966 for all copies at the Software Lib.)
$c (use "1", regardless how copies there are at the Software Library)
$d (do not use, even when there is accompanying material)
$l (Use "SOFTW")
$s (Use accession number, K— for Kinko's; W— for Wisc-Ware;C— for CONDUIT software)
$m (use "STACKS" initially for all items, use other material types once there are other sub-collections at the Software Library)
$b (wand the barcode)

A brief record example:

```
001    [PREFIX]99999999-D
100    10$aDavis, Stuart.
245    10$aPROSE student module.
250      $aVersion 1.2.
260      $aSanta Barbara, Calif. :$bKinko's Service Corp.,$c1986.
300      $a1 computer disk ;$c 3 1/2 in. +$e1 manual.
538      $aSystem requirements: Macintosh SE, Macintosh Plus;
         2 800K disk drives.
690     4$aEnglish language.
710    20$aKinko's Service Corporation.
700    10$aKaplan, Nancy.
700    10$aMartin, Joseph.
753      $aMacintosh.
966      $a1$lSOFTW$sK25$mSTACKS$b[barcode]
```

SS:cw

DIRECTORIES

UNIVERSITY OF HOUSTON

Computerize Your Research using CD-ROM databases

A number of journal indexes are now available in the library on CD-ROM (compact disc). A list is on the back of this flyer. Searching a CD-ROM database allows you to:

* Search complex topics such as company wellness programs and smoking cessation.
* Search multiple years of an index in a few minutes.
* Search for words or phrases in titles and abstracts, as well as subject headings, in most cases.

HOW MUCH DOES IT COST?

There is no charge for searching CD-ROM databases. There is a charge of 7 cents per page to print. Before printing you will need to purchase a Copicard from a machine in the Copy Center, Basement Red Wing. Downloading is free (bring your own MS-DOS formatted floppy disk).

WHERE ARE THE CD-ROM DATABASES LOCATED?

The CD-ROM terminals and assistant are located near the Reference/ Information Desk. You may search these databases whenever the Library is open. You may check-out the CD-ROM discs that you want to search from the CD-ROM assistant during the following hours:

Monday - Thursday	10 - 9
Friday	10 - 6
Saturday	12 - 4
Sunday	1 - 5

Other hours you may check-out the CD-ROM discs at the Reserves Desk, located at the end of the Circulation Desk.

The CD-ROMs are in very heavy demand. Use of a CD-ROM is limited to 30 minutes if someone else is waiting to use it. If you want to use a CD-ROM that is already in use, please sign up with the CD-ROM assistant. During the hours when there is no assistant, notify the person using the CD-ROM. That person is then entitled to a maximum of 30 additional minutes to complete his work and return the packet so that it may be checked out by the next user.

HELP!

Most of the databases have instructions on the screen and several include a tutorial. There is also an instructional flip-chart with each CD-ROM. An assistant is available to answer any questions during the hours listed above.

WHAT DATABASES ARE AVAILABLE ON CD-ROM?

ABI/Inform (1983 - present)
Indexes articles in business and management journals.
Printed equivalent: none

Compact Disclosure
Includes up-to-date information on public companies extracted from annual reports, 10-Ks, and related documents.

ERIC (1966 - present)
Indexes journal articles and unpublished reports in education and related fields such as communications.
Printed equivalent: Current Index to Journals in Education and Resources in Education, Index Table 5

Medline (1984 - present)
Indexes journal articles in medicine and related fields.
Printed equivalent: Index Medicus, Index Table 9

MLA Bibliography (1981 - present)
Indexes articles, books, dissertations, and other documents in the fields of literature, language, linguistics, and folklore. It provides worldwide coverage ranging from the Middle Ages to the present day.
Printed equivalent: MLA Bibliography, Index Table 13

Monthly Catalog (July 1976 - present)
Indexes publications of the United States government. Most of these publications are **not** included in the Library Catalog or the Serials List. This library receives a large percentage of these publications.
Printed equivalent: Monthly Catalog, Index Table 19

PAIS (1972 - present)
Indexes journal articles, books, and government documents covering public policy and social issues.
Printed equivalent: Public Affairs Information Service, Index Table 8

PsycLIT (1974 - present)
Indexes journal articles in the field of psychology and related fields.
Printed equivalent: Psychological Abstracts, Index Table 6

Science Citation Index (1988)
Indexes articles published in science and technology journals.
Printed equivalent: Science Citation Index, Index Table 2.

sociofile (1974 - present)
Indexes journals articles in the field of sociology.
Printed equivalent: Sociological Abstracts, Index Table 7

University of Houston Libraries
December 1988

SOFTWARE AVAILABLE FOR EVALUATION

AT

CCIS SOFTWARE INFORMATION CENTER

Busch Campus, Hill Center

Room 017

Faculty is encouraged to make appointments to review our software collection. For more information, please call Dr. Leny Struminger or Dr. May Lein Ho at 932-2427 or 932-2491.

77

Catalogs with detailed package descriptions are also available at the reference desks in each library. Interested faculty can request catalogs from the Software Information Center.

HOW TO USE THE SOFTWARE COLLECTION?

The Software Information Center is *open Monday through Friday from 9:00 am to 4:00 pm*. For other time, appointments can be made by calling 932-3129 or 932-2427. Designated microcomputers are provided for faculty to preview and evaluate the software. After evaluation, faculty can also check out selected software for a limited time.

HOW TO ORDER THE SOFTWARE?

The Center will provide faculty with necessary forms for *purchasing multiple copies of* software to be used in classroom instruction and research. A variety of class licenses is also available from the publishers. The staff will help faculty to implement their selected software. For detailed information, please contact Dr. Leny Struminger or Dr. May Lein Ho at *932-3129 or 932-2427*.

HOW ABOUT SOFTWARE DESIGNED BY RUTGERS FACULTY?

The Center will serve as a *depository for instructional software* created or adapted by Rutgers faculty. Faculty are encouraged to submit their software to be included in the collection. The center will also help faculty *market their computer programs through Wisc-Ware and Kinko's publishers*.

Table of Contents

Macintosh Courseware

	Package #
Anthropology	
MacPrimate.Vervet 3	001
Art	
Escher-Sketch 4	002
Authoring Tool	
Atlas and Overlay 5	003
CaseMaker 6	004
Drill 7	005
Guide Demonstration 8	006
Lesson Writer 9	007
Lesson Writer for Chemistry Lessons 10	008
MacLang 11	009
Matchmaker and Matchmaker Utility 12	010
Oyster 13	011
Biology	
Energetics and Metabolism 14	012
Evolution 15	013
Model Neuron 16	014
Plant Paint 17	015
Chemistry	
General Chemistry, Multiplan Templates 18	016
Huckel Molecular Orbitals 19	017
KSIMS 20	018
MacMendeleev 21	019
MacStereo (R/S) 22	020
Molecular Editor 23	021
Molecular Editor Demo 24	022
Molecular Weight Calculator 25	023
PeriChart 26	024
Computer Science	
Binary Trees 27	025
Event Tutor 28	026
Skel 29	026 (wi eve cut
ForthTalk Kernel 30	028
ForthTalk Library 31	029
Karel GENIE 32	030
Pascal Pointers 33	031
The SmallGol Compiler 34	032
Engineering	
FEMG: A Finite Element Mesh Generator 35	033
LogiMac 36	034
MacBode and MacLocus 37	035
Pipe Flow 38	036
Pipeline Design Kit 39	037
Signal Operations 40	038
TLS: Transmission Line Simulator 41	039
Transforms, Windows, Modulation 42	040
VLSI Magician 43	041
General Purpose	
Arabic Dates 44	042
Brainchild Grade 45	043
Drexel Plot 46	044
Glossary Maker 47	045
LaserTerminal 48	046
Learning Tool 49	047
Personal ResumeWriter 50	048
Reed Applications I 51	049
Reed Applications II 52	050
Tips Desk Accessory 53	051
Small Tips Desk Accessory	052
General Science	
Problem Solving Interpreter 54	053
Scientist's Spreadsheet 55	054
Geography	
Student Atlas–Counties 56	055
Student Atlas–U.S.A. 57	056
Student Atlas–World 58	057

(Continued)

Geology
GeoStructures 59 058

History
Treaty of Versailles 60 059
Vocabulary Drill 61 060
The Would-Be Gentleman 62 061

Law
Torts Exercises 63 062
XXXXXXXX 64
HARRIS VS. KLONDIKE KLIMBERS, INC 063

Literature
Aristotle's Greek Tragedy
Construction Kit 65 064
The TheaterGame 66 065

Mathematics
ALPAL 67 066
Big ALPAL 68 067
DEGraph 69 068
Differential Equations 70 069
FunPlot 71 070
FunPlot-3D 72 071
IVP 73 072
Linear Programming
by Fractions for the Macintosh ... 74 073
MacSimplex 75 074
Matrix Algebra 76 075
Not So Simple Arithmetic 77 076

Music
Appletones 78 077
Mozart 79 077
MacVoice 80 078

Nutrition
MacDiet—Student Version 81 079
MacDiet Expanded Database 82 080

Philosophy
BERTIE - II: A Symbolic
Logic Tutor 83 081
Tarski's World 84 082
Turing's World 85 083
Venn: A Syllogistic
Logic Tutor 86 084

Physics
DAFRAC—Diffraction
Calculation Tool 87 085
Heavenly Mac 88 086
Mechanics 89 087
Electromagnetism 90 088
Modern Physics 91 089
Optics Lab 92 090
Quantum Mechanics 93 091

Psychology
Hyper-Neuroanatomy 94 092
MacLaboratory for Psychology .. 95 093
MindLab 96 094
Neuroanatomy Foundations 97 095

Research
BiblioMania 98 096
BRS Simulator 99 097

Sociology
SocStat 100 098
SocStudy 101 098

Statistics
CLR Anova—student ver. 102 099
The Data Desk–student ver. 103 100
STAT HELPER I ..... 104 101 (ins
for v.)
STAT HELPER II ..... 105 102 (st
103 (in
for v.)
104 (st

Writing
Create 106 10
Idealiner Jr. 107 10
Spelling Champion 108 10
Tools for Writers 109 10

IBM COURSEWARE

Table of Contents

	Package #	
rts and Humanities		
dvertising Planner	1	016
uke Language Toolkit - Font Generation	2	041
GLOSSA - CAI for foreign language study	3	071
ILIM - Drills & Practice in Hebrew	4	068
SONNET - Intro to structural elements	5	070
tage Lighting	6	059
United States History	7	090
usiness/Economic		
QLPWIN - Quad. Linear Control Problems	8	072
VAL - Security Valuation Package	9	086
ngineering/Computer Science		
ASA - Assembly Systems Simulation	10	084
AFE - Connection Arrays	11	054
RITPATH - Critical Path Method Program	12	027
DC - Elementary Digital Computer	13	061
NS - Logic Network Systems (MS-DOS)	14	053
LNS - (MS-windows version)	15	078
icroprogrammed Computer Simulator	16	062
PC-Cube - Parallel Computing System	17	074
UFF - CAD for Microwave IC's	18	058
oil Mechanics Demonstrations	19	022
OLON - Transit Policy Evaluation	21	024
UNNEL - Limit Analysis of Tunnels	22	023
raphics/Image Processing		
SS MCL - Plotting Utilities	23	032
BM 5083 Driver	24	033
PCT - Prin. Component Transformation	25	082
LOT - IBM PC graphic plotting package	26	030
roGraph - prints EGA screen to Proprinter	27	020
DINetpl - IBM 7372 plotter device driver	28	001
DIPlot - Data Plotting routine	29	002
ealth Sciences		
nstitutional Patient Simulation	30	018
EPPER - Speech Processing	31	043
CB HG72 Diet Analysis System	32	057
SC PC-Pack - drug dosage research	33	039
ife Sciences		
ANTS - Ant Foraging Model	34	089
EPIGRAPH - Plant Disease Epidemic Sim.	35	079
# MEKA/MEKAEDIT - Plant Classification	36	055
* POPSIM - Plant Pop. Simulation Model	37	085
UnEven - Forest Predictions	38	015
Math/Statistics		
Data Transformations	39	036
* Differential Equations & Surfaces	40	088
DTREE -Decision Tree Analysis routine	41	026
FIT - non-linear Least Squares Fitting	42	004
LP - Linear Programming Problem Solver	43	025
NONMET II PLUS	44	063
PCFT - IBM PC Fourier Transform	45	029
* PREDCALC - Predicate Logic	46	076
Solver-Q - Nonlinear Equation Solver	47	044
Z-Score Practice Exercises	48	035
Philosophy		
Philo the Logician	49	014
Simulations for Philosophy	50	017
Physical Sciences		
APPLIED BASIC PHYSICS - Volume 1	51	067
* CIRC - Electronic Circuit Analysis	52	083
* CRYSTALLAB - X-Ray Diff. Simulation	53	080
* FLINCOR - Fluid Inclusion Isochore Calc.	54	081
# PCWAVE - E.M. Radiation Projects	55	031
* Physics Demos (MS-DOS)	56	077
* Physics Demos (MS-windows version)	57	075
SIMLAB - Statistical Simulations	58	038
SPEKL - Data Acquisition & Processing	59	034
VIVAS - Quantal Scattering Solver	60	042
Programming Tools		
Controlling the IBM 6 Pen Plotter	61	011
LogiTech Mouse Drivers	62	013
Professional Graphics Display Controller	63	010
PGC Libraries - IBM Professional Graphics	64	003
SCriptWriter - Authoring System	65	040
TurboGraphic Tools	66	056
TurboPascal EGA Drivers	67	012
YFORTH - editor, assembler & dictionary	68	028
Psychology		
Operant Conditioning	69	005
A Production System for Driving an Auto	70	037

* indicates a new addition in this edition. # indicates a new update in this edition. (Continued)

8i

Social Sciences

Fugawiland – an Archaeological Expedition ... 71 052
* MAPDATA – Data Mapping Application ... 72 087
Population Pyramids ... 73 050
* Texas Census Data System ... 74 069

Text Processing

GATOR – Text database system ... 75 066
* TRANSLIT – Text Filter ... 76 073

Package

Utilities

API – application program interface ... 77 0
ASK – batch file response handler ... 78 0
FONTGEN – display font generator ... 79 0
MV – moves files/sub-directories ... 80 0
PC/IP on NetBIOS – networking ... 81 0
RM – deletes files/sub-directories ... 82 0
UMCS – Bulletin board system ... 83 0

* indicates a new addition in this edition. # indicates a new update in this edition.

WHAT IS CCIS SOFTWARE INFORMATION CENTER?

To address the needs for academic software in higher education curriculum, a Software Information Center was recently established at CCIS, Busch Campus. The facility is located at the FACULTY/STAFF DEVELOPMENT CENTER, Room 017, Hill Center. At this center, ***instructional and research microcomputer software*** is available to faculty and graduate assistants for preview and evaluation purposes.

WHAT SOFTWARE IS AVAILABLE?

Rutgers has obtained a complete collection of software from WISC-WARE and from the Kinko's Academic Courseware Exchange. WISC-WARE is a consortium of colleges and universities, based at the University of Wisconsin-Madison, which distributes research and instructional software for IBM microcomputers. Kinko's software is basically for Macintosh computers. Both collections cover ***a wide variety of subject areas*** in many different disciplines, such as Arts & Humanities, Engineering, Health Sciences, and Mathematics. The Center is planning to obtain additional software from other sources based on faculty requests and suggestions.

WHY SHOULD THE FACULTY USE THIS CENTER?

At the Center, the faculty will be able to preview various software and to identify those useful for teaching or research. This software can be employed, alone or along with textbooks, to *teach new concepts, for homework assignments, or as remedial tools*. In addition, authoring tools and utility packages are available to facilitate computer use. Most of the software is menu driven and quite easy to use. Each computer program is accompanied by printed documentation. Using the Center's software may generate *new ideas for research and publications*.

HOW TO FIND OUT WHAT IS AVAILABLE?

Each software package is *listed in the IRIS system* which can be accessed from the terminals in any library on all campuses.

COURSES

NEW YORK UNIVERSITY

COVERING THE BASES: TECHNIQUES FOR DATA RESEARCH AND RETRIEVAL

A free, non-credit course for students and faculty to explore the availability, use and interpretation of print and computer-based data resources at NYU will be offered this winter. This 5-part course will begin Friday, February 5, 1988.

Topics to be included:

- accessing appropriate data resources
- print and online statistical resources
- creating and using local data files
- manipulation of data
- interpreting and displaying results

Course offers hands-on use of computerized data files such as the General Social Survey and the U.S. census, statistical software packages, and commercial online databases like Donnelley Demographics and Chase Econometrics.

Dates: Fridays, February 5 - March 4 (5 Sessions)
Time: 10 am - 12 noon
Place: Bobst Library, Avery Fisher Ctr., 2nd floor

For more information, contact the instructors:
Melanie Dodson, Bobst Library (998-2454)
George Sharrard, Academic Computing Facility (998-3402)

Enrollment is limited. Please register by returning the completed form to:
Melanie Dodson, Bobst Library, 70 Washington Square South, NY, NY 10012

NAME:
ADDRESS:
PHONE:
NYU ID NUMBER:

SCHOOL/PROGRAM:
DEPT/MAJOR:
Faculty: __ Graduate:__
Other: _____

Please indicate current research projects involving data resources and/or previous computing experience.

SEE ACRL Workbook CE 304 for curriculum materials related to this workshop.

UNIVERSITY OF FLORIDA

MRDF Orientation Course
UF Libraries
W. Covey, Instructor

Course Objectives

The purpose of this short course is to present a rudimentary overview of some of the concepts and techniques needed to work with machine readable data files, or MRDF. Participants will learn to evaluate MRDF for possible inclusion in the Libraries collections on the basis of documentation, suitability for our computer environment, and quality. In addition, participants will develop a better appreciation for the processes by which a dataset can be prepared to describe phenomena. Basic terminology relevant to discussing MRDF in an academic environment will be learned.

Organizational Information

The course will be conducted for four weeks, from 10 May 88 to 2 June 88. Meetings will be held from 1:00 to 1:50 every Tuesday, Wednesday, and Thursday. The instructor will be Bill Covey (Systems, Library West, 392-0796).

Work outside of class will consist primarily of construction and manipulation of two datasets, one constructed by each participant individually, and one supplied from the Libraries MRDF archive. Some small readings will also be supplied.

Periodically, evaluations will be made of participant progress. The results will be forwarded to the appropriate supervisors.

10-12 May

A. Overview, course organization, Data Matrix I

B. Measurement taxonomy, useful arithmetic

C. Arithmetic bases, coding schemes

17-19 May

A. Data organization, media types

B. Documentation, Data Matrix II

C. Sign-on, Data Entry and Editing

24-26 May

A. SAS I - basic data access and manipulation

B. Data Cleaning I - Frequencies & Crosstabulation

C. Statistics .02 - Measures of central tendency (mean, median, mode)

31 May - 2 June

A. Generic Computer Systems

B. Local Computer Systems

C. Summary

Prerequisites:

Scientific calculator (at least one memory, square roots, logs), Faculty Support Center sign-on, willingness to suspend disbelief

INSTRUCTIONAL GUIDES

Library Satellites Information Guide

USCInfo PERIODICAL INDEXES
Part I: General Description

What is USCInfo?

USCInfo is a collection of periodical index databases. The system is available at any of the libraries on campus and at the library satellites.

I. The Periodical Indexes

The Periodical Indexes provide the user with on-line access to selected indexing services. These services are machine-readable files similar to printed indexes like the Reader's Guide to Periodical Literature, Business Periodicals Index and Public Affairs Information Service (P.A.I.S.). Currently, the library provides access to five online indexing services (listed below). These indexes provide citations from over 1,000 periodicals, newspapers, and books published in the United States and Canada. There are over two million citations in these databases and they are updated monthly.

Magazine Index

Magazine Index is a comprehensive index of the most frequently cited journals from the United States and Canada. Approximately 400 magazines are indexed covering general interest as well as highly specialized publications. Indexing coverage begins with 1983 with monthly updates. Subject coverage includes:

current events	social sciences	popular opinion
library science	business & economics	science & technology
art & architecture	health & medicine	travel, sports & leisure
education	cultural events	consumer affairs
performing arts	literature	

Management Contents

An index including abstracts covering the basic functions of management and administration. Indexing begins with 1986 with monthly updates for 127 key management journals, newsletters and tabloids. Subject coverage includes:

management	decision science	economics & finance
accounting	auditing	personnel & benefits
training	government	public administration
industrial relations	labor relations	production
government regulations & legislation		advertising & marketing
distribution	sales	

National Newspaper Index

A comprehensive index of the five most frequently cited newspapers in the United States. The newspapers indexed are the Christian Science Monitor, the Los Angeles Times, the New York Times, the Washington Post, and the Wall Street Journal. Indexing coverage begins in 1981.

Computer Database

An index with abstracts covering major computer industry journals. Articles indexed cover all types of systems from microcomputers to mainframes. Eighty-five computer journals are indexed. Indexing begins in 1986 with monthly updates. Subject coverage includes:

computer systems	software & hardware	telecommunications
product announcements	design & development	information management
high-tech companies	computer services	programs & programming
MIS (Management Information Systems)		product evaluations
computer applications		

Trade and Industry Index

A comprehensive index of over 800 business, trade and industry magazines, journals, and tabloids covering a wide range of industries. Indexing begins with 1982 and is updated monthly. Subject coverage includes:

advertising	agriculture & forestry	oil & gas
construction	transportation	design & manufacturing
food	public utilities	wholesale & retail trade
finance	insurance	real estate
banking	securities	commodities
taxation		

COMING SOON:

II. **HOMER** - The USC Online Catalog: The CATG database is the USC Library Catalog. It will soon be available on the USC Information System. Please use the HOMER system to access the Catalog at this time.

III. **USC Library Union List of Serials:** The ULIS database is the USC Library Union List of Serials. It will soon be available on the USC Information System. Please consult a print version of the Union List at any campus library reference desk or at the library satellite locations.

IV. **USC Campus Information System:** The USC Campus Information System will include a calendar of events, phone directory of faculty, staff and students, and hours of operation and location of campus facilities and services.

For more information about USCInfo, take a look at the following handouts (available at any Library Satellite):

Part II: LogOn/Off Procedures
Part III: Quick Guide to Operation
Part IV: Search Techniques
Part V: Troubleshooting

LSIG-001a LSW 08/88 (rev. RMP 08/88 [OWS])

USCInfo PERIODICAL INDEXES
Part II: LogOn/Off Procedures

	Satellite Locations	Dial-Up (Modem)	Micom Connections
LogOn	1) From the main network menu choose "2 Library Resources". 2) When the next menu appears, "2 USCInfo Databases" should be chosen. It will now take a short time log in.	The USCInfo disk is bootable. You can start your computer with it and be automatically logged into USCInfo. To do this, place the disk in your A drive, turn the computer on and in a few moments you will be in the system. If your computer is on, type (A:) AUTOEXEC and you will be logged in.	1) At the "Which system?" prompt enter USCVM. 2) Next you will be asked for your terminal emulation. Supply the correct information (normally VT100). 3) You will next be asked for your ID, then password. Enter them and you will be in USCInfo.

Choose search level 1. NOTE: If you find yourself in the more advanced level (native mode) by mistake and would like to go back to the menu-driven mode, you must exit the system and start over. Enter ..o (period, period, letter "o"), then wait until the "Which system?" prompt appears. Follow the LogOff procedures, detailed below, and try again. See Parts III and IV for details on searching.

LogOff	1) Type =x and press ENTER. 2) Wait until the "Which system?" prompt appears. At this point, hold down the ALT key and press the X key. 3) The system will query, "EXIT TO DOS? (Y/N)". Type Y and the main network screen will appear.	a) Type =x and press ENTER. b) At the next screen, type X and press ENTER again. c) Wait until the "which system?" prompt appears. At this point, hold down the ALT key and press the X key. d) The system will query, "EXIT TO DOS? (Y/N)". Type Y and the A> prompt will appear.	1) Type =x and press ENTER. 2) Wait until the "Which system?" prompt appears. At this point, hold down the ALT key and press the X key. 3) The system will query, "EXIT to DOS? (Y/N)". Type Y and you will return to the DOS prompt.

For more information about USCInfo, take a look at the following handouts (available at any Library Satellite):

Part I: General Description
Part III: Quick Guide to Operation
Part IV: Search Techniques
Part V: Troubleshooting

USCInfo PERIODICAL INDEXES
Part III: Quick Guide to Operation

Command Summary

=S - to begin a new search.
=C - to change to a different database.
=H - to view online explanation of commands.
=N - to read news of the LIS system.
=O - to set searching options.
=X - to log off of system; additional steps are required.

Boolean Operators

AND - retrieves citations with all keywords present; default (assumed by system) unless changed on the options menu.
OR - retrieves citations with any of the keywords.
NOT - retrieves citations with first, but not second keyword.
XOR - retrieves citations with either keyword, but not both.

Examples: "constitution freedom speech" retrieves only those citation that have all three keywords present; "constitution or freedom" retrieves citations with either or both of the keywords; "constitution not freedom" retrieves those with constitution as long as they do not also contain freedom; "constitution xor freedom" retrieves those with either, but not both.

Positional Operators - must be used with a component designator

SAME - only those citations with your keywords in the same component are retrieved.
ADJ - restricts the search to citations with keywords next to each other, in the order entered, in the same component.
WITH - restricts the search to citations in which the keywords appear in the same sentence.

Examples: "constitution same freedom.de." retrieves citations with both keywords in the same component; "constitutional adj freedom.ti." retrieves titles with the phrase "constitutional freedom" in them; "constitution with freedom.ab." retrieves citations where both keywords appear in the same sentence in the abstract.

Component Designation

Any search may be restricted to a particular component by adding its abbreviation to the end of your keyword string, e.g., constitution.ti. (for title component) or freedom same speech.de. (for descriptors component). When using positional operators (see above), the component designator should always be included. Designating which component is to be searched is particularly useful for limiting the number of items you find (e.g. when your search finds more than 200 items).

For more information of using "component designation" and a list of the components, take a look at Part IV: Search Techniques.

Downloading & Printing Citations

If you are using Procomm software, press ALT-G to save each screen to a file (Satellite Locations save to the B: drive as BRS.SCH). This file can be printed directly, or you may use a word-processor to rearrange the file to your liking. Ask the consultant at the Satellite for assistance. NOTE: If you are using other communications software, please check your manual for procedures.

To retrieve periodical issues, consult the USC Union List of Serials, available at any USC Library and the Library Satellites.

For more information about USCInfo, take a look at the following handouts (available at any Library Satellite):

Part I: General Description
Part II: LogOn/Off Procedures
Part IV: Search Techniques
Part V: Troubleshooting

USCInfo PERIODICAL INDEXES
Part IV: Search Techniques

The USCInfo Periodical Indexes provide on-line access to selected periodical indexes. These services are similar to printed indexes like the Reader's Guide to Periodical Literature, Business Periodicals Index or Public Affairs Information Service (P.A.I.S.). The system is available at any of the libraries on campus and at the library satellites.

The system is menu-based. You choose options from screen menus and enter 1 to 3 character commands. For further help, consult this handout or ask the library staff. **NOTE:** The ALT-F10 Help key, shown at the bottom of the screen, does not give help for the USCInfo system, but rather for the communications system supporting it.

Once logged on, the following screen will appear:

```
**********************************************************************
USCINFO                                                      08/17/88
Please select one of the following:
   1 - Search a database at beginner level
   2 - Search a database using BRS/Native Mode Interface
   3 - Use the USC Campus Information System
   X - Exit USC Library Information System
Enter your selection; H for help (H):
**********************************************************************
```

Choose search level 1 to see the next screen:

NOTE: If you find yourself in the more advanced level (**native mode**) and would like to go back to the menu mode, you must exit the system and start over. Enter ..o (period, period, letter "o"), then wait until the "**Which system?**" prompt appears. Follow the **LogOff** procedures, detailed below, and try again.

```
**********************************************************************
MAIN menu                                          Wed Aug 17 1988
Select one of the following functions by typing a letter and
pressing the RETURN key:
   S begin searching a database
   C select a database to search
   H get help
   N for news about LIS
   O set search options
   Q quit and return to previous menu
Enter selection:
**********************************************************************
```

The MAIN menu shows the available commands:

S - Search a database This allows a user to choose a database to search and then to input words to locate citations.

C - Change database	This allows the user to change databases for searching
H - Online help	This allows the user to read a description of available commands and other system information.
N - News	To read news about updates and changes to the system.
O - Options	To change the assumed Boolean operator for searching; the default operator is "and"
X - Logoff	This logs the user off the system. See additional LogOff instructions at the end of this guide.

To continue, type s and press the return key. The following screen appears:

```
**********************************************************************
USCINFO                                                Thu Feb 18 1988
CHANGE database menu

Enter =C to set databases, =H for help, =O to set options, =X to
exit, or enter a number to select one of the listed databases:

1 MAGX - Magazine Article Index          5 CATG - Card Catalog
2 MCON - Management Contents             6 ULIS - Union List of
3 TRIX - Trade and Industry Index                 Serials
4 COMP - Computer Databas                7 NNIX - National News-
                                                  paper Index

Your choice?
**********************************************************************
```

In order to continue searching, you must select a database. To select a database, type the number and then press the return key. If you wish to see a description of a database, enter "=h" and then choose a database description. If you select a database (for example, Magazine Index), you will see:

```
**********************************************************************
USCINFO                Database: MAGX                  Thu Feb 18 1988
You are now signed on to the MAGX database.
Press RETURN to continue ..
**********************************************************************
```

Press return to go to the following screen:

```
**********************************************************************
USCINFO                  Database: MAGX              Thu Feb 18 1988
SEARCH menu

Enter =C to set databases, =H for help, =O to set options, =X to
exit, enter Q to go back to the previous menu or enter a list of
words to search.

No searches outstanding.

Type search words and press RETURN :
**********************************************************************
```

When you type the keywords you wish to search, the system assumes **AND** between each word. You must specify **OR, NOT** or **XOR** or you can change the default operator. For details, see the section, **Boolean Searching**, on page 5.

As an example, assume you entered "**joint ventures japan**". You would see the following screen:

```
**********************************************************************
USCINFO                  Database: MAGX              Thu Feb 18 1988
SEARCH menu

Enter =C to set databases, =H for help, =O to set options, =X to
exit, enter Q to go back to the previous menu, press RETURN to
browse citations, or type a list of words to search.

Search results:
1: VENTURES AND JOINT AND JAPAN found 340 times in 50 citation(s).

Enter search words or press RETURN to browse results:
**********************************************************************
```

To see the citations which have resulted from your search, press ENTER. The citations appear in reverse date order, that is, most recent date first. Your search words will be highlighted.

To continue to display citations press the return key. To quit viewing citations, type Q and press return.

Downloading and Printing a Citation

If you are searching USCInfo using a copy of Procomm supplied by the Library Satellites, follow the instructions below. If you are using another communications software package or are using a copy of Procomm that was not supplied by the library, follow the directions in your manual for doing a "screen dump" to a file.

1. To download a citation, place a formatted disk in drive B of the computer, hold down the **Alt** key and press the G key. This places a copy of the screen on the disk in drive B in a file called BRS.SCH. After you have logged off, use DOS or any word processing program to print out the results of your search.

2. If you want to keep the file after you have printed it, you should use your word-processor (or DOS if you are **not** doing this at a Library Satellite) to rename the file. If you do not do this and use the same disk again, the citations will be overwritten with new text the next time citations are copied to disk.

To find periodical locations, consult the USC Union List of Serials, available at any USC Library and in the Satellites. The Union List provides an alphabetical listing of all magazines, journals, newspapers, and government documents held by the USC Libraries.

Constructing a Search

Developing a search for on-line indexing and abstracting services differs from searching the on-line catalog. With the on-line indexing and abstracting services you do NOT have to specify author, title or subject, although you may. Instead, you search the entire contents of the record by keyword. -

Each database contains thousands of citations (records) for periodical articles. Each citation consists of separate components (fields) such as title, author, name of periodical, and subject descriptors. Each component breaks into sentences and each sentence into words. Using Boolean search techniques, words in a citation, component or sentence can be searched.

The simplest way to construct a search is to type a string of keywords (e.g. "constitution freedom press") which represent a topic that interests you. Do not include prepositions (in, with), conjunctions (and, or, but), subjunctives (while, which, what), or punctuation (except for component designation). These words and punctuation are not included in the indexing and their use results in your search retrieving no citations. The system automatically assumes that "AND" is between your keywords. For information on using other operators or relationships between terms, see **Boolean Searching** below.

If you do not know precise terms for your search, use more general terms to construct a search. Look at the results, and then, from the titles and descriptors retrieved, select more specific terms to compose another search.

The system will not display citations from a set larger than 200 records. If you have a large set, you must narrow your search by using more keywords. For example, if search #1 was "**joint ventures**" and there were 867 entries, you could then enter "1 japan". The system combines set #1 with the new term to come up with 50 entries.

Boolean Searching

Boolean operators specify the relationship of words to each other. USCInfo uses four operators: and, or, xor, and not.

AND - The system assumes **AND** between keywords, and therefore searches only for those citations containing all of the keywords entered.

ex. **constitution freedom speech** retrieves citations containing:
1. constitution and freedom and speech

Press return to go to the following screen:

```
**********************************************************************
USCINFO                  Database: MAGX              Thu Feb 18 1988
SEARCH menu

Enter =C to set databases, =H for help, =O to set options, =X to
exit, enter Q to go back to the previous menu or enter a list of
words to search.

No searches outstanding.

Type search words and press RETURN :
**********************************************************************
```

When you type the keywords you wish to search, the system assumes **AND** between each word. You must specify **OR**, **NOT** or **XOR** or you can change the default operator. For details, see the section, **Boolean Searching**, on page 5.

As an example, assume you entered **'joint ventures japan'**. You would see the following screen:

```
**********************************************************************
USCINFO                  Database: MAGX              Thu Feb 18 1988
SEARCH menu

Enter =C to set databases, =H for help, =O to set options, =X to
exit, enter Q to go back to the previous menu, press RETURN to
browse citations, or type a list of words to search.

Search results:
1: VENTURES AND JOINT AND JAPAN found 340 times in 50 citation(s).

Enter search words or press RETURN to browse results:
**********************************************************************
```

To see the citations which have resulted from your search, press ENTER. The citations appear in reverse date order, that is, most recent date first. Your search words will be highlighted.

To continue to display citations press the return key. To quit viewing citations, type **Q** and press return.

Downloading and Printing a Citation

If you are searching USCInfo using a copy of Procomm supplied by the Library Satellites, follow the instructions below. If you are using another communications software package or are using a copy of Procomm that was not supplied by the library, follow the directions in your manual for doing a 'screen dump' to a file.

1. To download a citation, place a formatted disk in drive B of the computer, hold down the **Alt** key and press the **G** key. This places a copy of the screen on the disk in drive B in a file called BRS.SCH. After you have logged off, use DOS or any word processing program to print out the results of your search.

OR - By specifying OR between words in a search, the system looks for citations containing one or more of the keywords.

ex. constitution or freedom retrieves citations containing:
1. constitution
2. freedom
3. constitution and freedom

XOR - the XOR operator retrieves either keyword but not both keywords.

ex. constitution xor freedom retrieves citations containing:
1. constitution
2. freedom
Citations with both freedom and constitution are not retrieved.

NOT - used to retrieve citations with one keyword but NOT another.

ex. constitution not freedom retrieves citations containing:
1. constitution but not freedom
Constitution and freedom are not retrieved.

When no other Boolean operator has been specified, the system assumes AND between the terms you enter. If you plan to do several searches calling for the same operator, it may be useful to change the default. This assumption, or default, can be changed using the OPTIONS menu.

Positional Operators

An excellent way to narrow your search is use of positional operators with a specific component in the record. These operators specify the relationship of words within a component.

The component abbreviation (ti, au, de) should be used as part of the search. USCInfo uses three positional operators: SAME, ADJ, and WITH.

SAME Use when you wish the keywords to appear in the same component.

ex. constitution same freedom.de. retrieves descriptor components with both 'constitution' and 'freedom' in them.

ADJ Use when you wish the keywords to occur next to each other, in the order entered, in the same component.

ex. constitutional adj freedom.ti. retrieves titles with the phrase constitutional freedom.

WITH Use when you wish the keywords to appear in the same sentence in the component, as in an abstract.

ex. constitutional with freedom.ab. retrieves abstract components in which 'constitution' and 'freedom' appear in the same sentence.

Component Designation

Article citations are made up of components, each of which hold a particular type of information. The meaning of the components and their two-letter abbreviations follows:

DO	BRS Assigned Accession Number
AU	Author (100)
TI	Title of Article
TT	Title (Poem or Story)
JN	Journal Name
PF	Fixed Publication Date
DT	Displayable Publication Date
YR	Publication Year
VL	Volume Number
IS	Issue Number
PG	Page Number
AB	Abstract
DE	Descriptors (1000)
BI	Biographee (100)
NA	Named Person (100)
RA	Revieved Author's Name (100)
GC	Geographic Code (100)
GL	Geographic Location (100)
IL	Illustrations (100)
CN	Company Name (100)
PN	Product Name (1000)

SHRT: AU, TT, JN, DT, YR, VL, IS, PG, DE

Any search can be restricted to a particular component by use of .(component abbreviation). at the end of a keyword string. By using "component designation" you will limit the number of items you will retrieve. This is useful when your original search finds more than 200 items and you need to bring the number down to browse the set.

ex. constitution.ti.

or

freedom same speech.de.

For more information about USCInfo, take a look at the following handouts (available at any Library Satellite):

Part I: General Description
Part II: LogOn/Off Procedures
Part III: Quick Guide to Operation
Part V: Troubleshooting

USCInfo PERIODICAL INDEXES
Part V: Troubleshooting

Problem: The screen shows the number to dial, but no connection is made. The modem is silent.

Solution: The modem may be set to the wrong port. First press the Esc key to terminate the automatic logon. Then hold down the Alt key and press x and exit to DOS. Then go to the section below on editing the brsvm.cmd file.

Problem: Connection is made, but ProComm redials anyway.

Solution: Due to telecommunications delays, the connection takes longer than the ProComm settings allow. First press the Esc key to terminate the automatic logon. Then hold down the Alt key and press x and exit to DOS. Then go to the section below on editing the brsvm.cmd file.

Problem: Connection is made, but the screen display is garbled.

Solution: You may have gotten a bad phone line. First press the Esc key to terminate the automatic logon. Then hold down the Alt key and press x and exit to DOS. Then type autoexec to try the automatic logon again. Alternatively, use the manual logon procedure. If you have trouble repeatedly, please call (213) 743-3455 for assistance.

Problem: ProComm dials the wrong number.

Solution: First press the Esc key to terminate the automatic logon. Then hold down the Alt key and press x and exit to DOS. Then go to the section below on editing the brsvm.cmd file.

Problem: Automatic logon appears to have stopped midway.

Solution: Wait at least 60 seconds before attempting anything.

If the logon has really halted midway, press the Esc key to terminate the automatic logon. Then hold down the Alt key and press X and exit to DOS. Type autoexec to begin again. If it again halts, please call (213) 743-3455 for assistance.

Problem: While in the Periodical Indexes, the system suddenly "throws you out" or gives you a message you don't understand.

Solution: First hold down the Ctrl key and press g to try to begin again from where you left off. If this doesn't work, type **logoff** and press return. This may get you out to the **Which system?** prompt, from which you can exit to DOS and begin again.

If neither of these works, hold down the **ALT** key and press the F7 key, and then release. Repeat twice more. This should get you to the **Which system?** prompt, from which you can exit to DOS.

If none of this works, turn off your computer. You will be automatically logged off after thirty minutes.

Problem: The "screen dump" (downloading), saves to the wrong drive, or disappears.

Solution: The ProComm software is set to download onto the b: drive, into a file named "brs.sch". If this is not the case, or you want the file to go elsewhere:

1) Boot the computer with a DOS disk (not the USCInfo disk).

2) Remove boot disk and place USCInfo disk in the A: drive. Type "procomm" and press return.

3) When ProComm screen appears, hold down the Alt key and press s. A **SETUP MENU** appears. Select **4) GENERAL SETUP.**

4) From the next menu, select **3) Default log file.** Type b:log.brs and press return. Next select **4) Screen dump file.** Type b:brs.sch and press return.

5) When finished, press the Esc key. At the next menu, type s, press return, and then press the Esc key again.

6) Leave ProComm by holding down the Alt key and pressing x and exiting to DOS.

Editing the Automatic Logon File

One of the difficulties with sending out disks to multiple locations is that sometimes the disk may not match your system set-up exactly. If this is the case, you can edit the Automatic Logon File, called **BRSVM.CMD** and the startup file called **PROFILE.CMD**, so that it will operate. If you do not wish to attempt this unaided, Lucy Siefert Wegner (213) 743-3455 can provide help.

The disks have been configured with the following settings: 1200 baud, 8 data bits, 1 stop bit, no parity, port com1, vt100 emulation.

Below is a listing of both the **PROFILE.CMD** and **BRSVM.CMD** files. You can see them onscreen by using the **MED** editing program included on your USCInfo disk. This program is a very simple text editor that works similarly to a word processor.

From the A> prompt, type **med brsvm.cmd**. Any changes you enter and save will change the way the automatic logon operates. See the listing below for where to make various changes.

```
[PROFILE.CMD]
     clear
     set port com1                    [If the port is incorrect, change "com1"
                                                to "com2".]
     set duplex full
     set parity none
     set databits 8
     set stopbits 1
     set baud 1200                    [If you want to try another baud rate,
                                       change it here. Be sure to change the
                                        phone number in the BRSVM.CMD file,
                                                listed below.]
     emulate vt100


[BRSVM.CMD]
     clear

     dial:
       locate 24 1
       transmit "ATDT 743-8753!"      [If the phone number is incorrect, change
                                                it here.]
       waitfor "connect"
       if waitfor
         goto connect
       else
         transmit "ATH0!"
         clear
         locate 11 30
         message "Re-dialing USC network..."
         goto dial
       endif

     connect:
       locate 8 25
       message "Automatic LogOn -- please wait."
       alarm 1
       transmit "!"
       waitfor "Which system?" 3
       if waitfor
         clear
         locate 8 30
         message "Logging in..."
         goto uscvm
       else
         goto connect
       endif

     uscvm:
       locate 10 1
       transmit "USCVM!"

     login:
       transmit "!"
       waitfor "ENTER TERMINAL TYPE:" 3
       if waitfor
```

192

```
    transmit "vt100!"
    pause 1
    transmit "!"
    pause 1
    transmit "1 XXXXXX!"        [This is where your logon id appears.]
    pause 1
    transmit "XXXXXX!"          [This is where the password appears.]
  else
    goto login
  endif

exit
```

To save your changes, press F10 twice. To exit and not save changes, press F10, and then F8 twice. If you have attempted to correct your logon file and instead made things worse, please call (213) 743-3455 for assistance.

Manual Procedures

Logging On

1) Boot your computer with DOS system.

2) Place the USCInfo disk in the A: drive.

3) Type procomm and press return.

4) When the procomm screen appears, hold down the Alt key and press d. The dialing screen appears.

5) Type m and then type the USC computer number and press return:

300 baud -	(213) 743-7646
1200 baud -	8753
2400 baud -	5000

6) The system will dial and the Which system? prompt appears.

7) Type uscvm and press return.

8) System will prompt, ENTER TERMINAL TYPE:. Type vt100 and press return.

9) The USCVM logo screen appears; press return.

10) System prompts Enter one of these commands: [etc]. Type logon and then your logon id (this is written on your disk). Press return.

11) System then asks ENTER PASSWORD...:. Type doheny and press return.

12) The USCInfo screen appears as detailed on page two of the documentation. Procedure from this point is the same as with the automatic logon procedure.

Logging Off The manual logoff procedure is the same as that used with the automatic logon.

1) Type =q and press return.

2) Wait until the which system? prompt appears.

3) Hold down ALT key and press x key.

4) System queries,.. 'EXIT TO DOS? (Y/N)".

5) Type y and the DOS prompt appears.

For more information about USCInfo, take a look at the following handouts (available at any Library Satellite):

Part I: General Description
Part II: LogOn/Off Procedures
Part III: Quick Guide to Operation
Part IV: Search Techniques

Searching USCInfo

To begin, click on Campus Info. Next, click on **Go Ahead**, then click on the "**LogOn USCinfo**" button at bottom of the screen. The Mac will automatically log you onto the program. Wait for this screen to appear:

```
Please select one of the following:

    1 - Search a database at beginner level
    2 - Search a database using BRS/Native Mode Interface
    3 - Use the USC Campus Information System
    Q - Exit USC Library Information System

Enter your selection; H for help (H):
```

Type in a 1 then press **return**. A screen saying BRS Technologies will appear; press the **ENTER** key (lower ht corner of keyboard) to continue. The Main menu will appear. Type C then press **return** to select a database to rch:

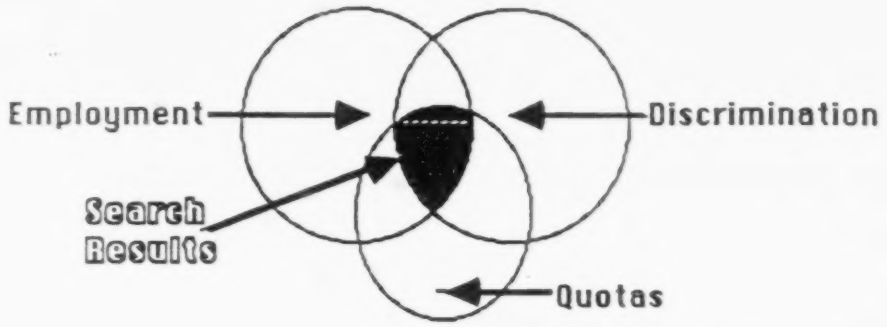
```
MAIN menu

Select one of the following functions by typing a letter and pressing
the RETURN key:

    S begin searching a database
    C select a database to search
    H get help
    N for news about LIS
    O set search options
    Q quit and return to previous menu

Enter selection:
```

Type in the number (1-4) of the database to be searched then press **return.**

```
CHANGE database menu

Enter =C to set databases, =H for help, =O to set options, =Q to quit,
press RETURN to go back to the previous menu, or enter a number to
select one of the listed databases:

 1 MAGX - Magazine Article Index      5 CATG - Card Catalog
 2 MCON - Management Contents         6 ULIS - Union List of Serials
 3 TRIX - Trade and Industry Index
 4 COMP - Computer Database

Your choice? 1
```

The Main Menu will reappear. Type an S then press **return** to begin a search. Type in the 2 or 3 words which

best describe your topic then press return. The computer will retrieve only records in which ALL your terms appear.



5. The computer responds by listing the number of times each term appears separately. On the last line terms are combined to show the number of records in which ALL terms appear. To see your results, type in the search number (e.g. 1) and press return. To enter a new search, type in your key terms and press return.

6. On the next screen, type in a B to browse the citations on your topic. The computer displays one record at a time; the most recent article will appear first. Press return to see the next record. Type a Q to stop browsing.
Sample record:

TI = Title of article
JN = Name of journal
DT = Date
YR = Year
VL = Volume number
PG = Page number(s)

Use the descriptors to get ideas for other searches.

```
Browsing citation number 1 of 21...

 Document 1
DO  000766928
Author
1   Pols, Cynthia
TI  Supreme Court ruling on racial quotas still not definitive.
JN  Nation's Cities Weekly
DT  March 2
YR  1987
VL  v10
PG  p11(1)
Descriptors
1   United States. Supreme Court cases
2   Racial quotas legal aspects
3   Discrimination in employment cases
4   Affirmative action programs legal aspects

Enter Q to quit browsing, press RETURN to continue:
```

7. To print citations, position the arrow over **Photo** (in lower left corner of screen) and click. Before the first photo, you will be asked to name your photo file; use your name. All your photos will be added to this file. Repeat for each record to be copied. After exiting USCInfo use Microsoft Word to print the file (see #9 below). To see a full list of commands type =B.

8. To exit, use mouse to position arrow on **Quit** (lower right corner of screen) and click once. System will automatically logout and return to the desktop.

9. To print, click on **Word.**Use mouse to position arrow on **File** menu; hold down mouse button and pull down menu until **Open** is highlighted; highlight your file name & click on **Open;** pull down **File** menu to highlight **Print;** click on **Faster** then **OK** to print. When finished, pull down **File** menu and highlight **Quit.** To eject the diskette, press **command (apple) key, shift, 1.**

ABI/INFORM

The ABI/Inform Business Database consists of abstracts and indexes to articles contained in business and management journals.

To print your results, you will need to purchase a Copicard in the Library Copy Center, Basement Red Wing. Printing costs 7¢ per page.

For additional help, consult the monitor near the CD-ROM Workstations or the staff at the Reference/Information Desk.

When you are finished searching ABI/INFORM, you must exit from the program before removing the compact disc. For instructions, see the page entitled QUITTING.

UNIVERSITY OF HOUSTON LIBRARIES--February 1989

This is a flip chart.

BEGINNING A SEARCH

1. Insert the computer disk into the drive and latch the cover.
2. Insert the CD-ROM disc into the CD-ROM player and press the button to close the drawer.
3. Follow the instructions on the screen.

Begin your search with the **Main Menu**. If the **Main Menu** is not on the screen, press the **F10** key.

The **Main Menu** offers several options. Use the ↑↓ arrow keys on the right side of the keyboard to move the highlight bar to the option you want. After highlighting the option you want, press the ↵ key.

To start a search, highlight **Search the ABI/INFORM Database**, then press the ↵ key.

For a description of what is included in ABI/Inform, highlight **Introduction to ABI/INFORM**, then press the ↵ key.

For a tutorial of the search software, highlight **How to Use ABI/INFORM**, then press the ↵ key.

Help

Help is always available after you start a search. Press the **F1** key for help at any point in your search.

The bottom line of the screen displays user instructions and the active function keys. To scroll through all of the **commands** available, press the **F2** key.

BEGINNING A SEARCH

Begin your search on the screen that has the **Search terms:** prompt. You will automatically reach this screen when you choose the **Search the ABI/INFORM Database** option from the **Main Menu**. To get the **Main Menu**, press the **F10** key.

Your search may consist of a simple one-word term or a phrase, or it may be a more complex combination of ideas. For complex searches, begin with a simple search of one of your concepts, then combine it using set numbers or additional terms as illustrated below. A list of suggested subject terms can be found in the Controlled Vocabulary section of the Search Inform guide near the workstation.

Simple/Short Searches

Type in a single **word** or short **phrase** to search, then press the ↵ key.

EXAMPLES: trademarks
advertising campaigns
centralized management information systems

SEARCHING

COMBINING SEARCH TERMS

Combine terms with Boolean operators to narrow or broaden a search.

Operator:	Item will be retrieved if:
and	Both terms are present
or	Either term may be present
and not	First term but not second present

EXAMPLES: labor relations and robotics
microcomputer or PC
Exxon and not oil

Combining Terms Using Search Sets

Search sets can be reused and combined by using the set number enclosed in square brackets []. After doing your initial simple search, press the **F3** key. Your first search will be given a set number in brackets that can be used to combine with other terms or other set numbers.

EXAMPLES: [1] or [3]
[2] and underwriting

Complex Search Expressions

You may use parentheses () to group terms

EXAMPLES: oil spills and (litigation or tort)
(compact disc or videodisc) and retail

Truncation, the use of proximity operators, and methods for restricting your search to a specific field are discussed on the **ADVANCED SEARCHING** page of this chart.

COMBINING SEARCH TERMS

ADVANCED SEARCHING

Truncation

End a term with ? to search all possible endings.

EXAMPLE: person? will search for person, personal, personnel, etc.

Proximity Operators

Proximity operators specify relationships between terms.

Operator:	Relationship:
w/#	First term within # words of second
pre/#	First term precedes second by exactly # words
w/seg	Both terms present in same database field

EXAMPLES: communications w/7 satellite
federal pre/2 commission
jan w/seg 1987

Searching by Field

To search for a term only within a specific database field, type the field name (or abbreviation); then enclose the search term in parentheses ().

Most commonly used field names and abbreviations:

Title	TI	Duns	DU
Journal	JO	Terms	TE
Jrnlcode	JR	Codes	CD
Date	DA	Abstract	AB
Company	CO		

EXAMPLES: company (Sears)
terms (direct mail)
CO (Citicorp)
JO (Forbes)

Terms used in the TE field are listed in the Controlled Vocabulary section of the Search Inform guide near the terminal.

SEARCHING TIPS

LIMIT YOUR SEARCH TO ARTICLES THAT HAVE BIBLIOGRAPHIES. All articles that have bibliographies have the word "References" at the end of the abstract. By searching for the word "References" in the abstract field, you will focus on these more scholarly/research articles.

EXAMPLE

[1] fringe benefit plans
[2] ab(references)
[3] [1] and [2]

FIND CASE STUDIES. Simply search for the term "Case Studies" in the terms field and add it to your strategy.

EXAMPLE

[1] operations research
[2] te(case studies)
[3] [1] and [2]

RESTRICT YOUR SEARCH TO ARTICLES ON THREE OR MORE COMPANIES. This is one easy way to get articles that compare several companies' policies, responses to a problem, etc.--useful for industry information. To get these articles, add the word "Manycompanies" to your strategy.

EXAMPLE

[1] personnel policies
[2] manycompanies
[3] [1] and [2]

Also available: Manycountries, Manyindustries, Manypeople, Manyproducts. Try them!

USE THE THESAURUS. Although you an make up phrases and search terms of your choosing, the Search Inform thesaurus will help you to select the best search terms. Search Inform can be found on Index Table 4--look for a small, gray looseleaf binder.

SEARCHING TIPS

112

VIEWING SEARCH RESULTS

Your search strategy will be displayed. You may choose to:

A. See a list of titles from which you can choose the citations you wish to read.

B. Begin by viewing the first citation with the abstract and page through your results.

TO PRINT OR DOWNLOAD YOUR RESULTS, MARK THE DESIRED CITATIONS AS EXPLAINED BELOW, THEN SEE THE NEXT PAGE OF THIS CHART.

A. Title List

To list titles of retrieved documents, press the ↵ key. Choose the titles you wish to view one at a time by moving the highlight bar to the desired title. To move through the title list use ↑↓, **PgDn**, and **PgUp**. To see a citation and abstract, highlight the title and press the ↵ key.

To return to the **title list** and select another document, press the **Esc** key.

If you intend to print or download the citation after viewing your results, press the **F9** key or the **spacebar** to mark the title as a desired citation before moving on to your next choice. If you change your mind, press either key again to unmark the choice.

B. Citation Display

To see the first citation with the abstract, press the **F7** key. To view the **next** and **previous** document in the retrieved set, use the **+** and **-** keys. These keys are on the far right of the keyboard.

To **scroll** through a single citation record (in order to read the bottom of the abstract), use **PgUp** and **PgDn**. These keys are on the right of the keyboard.

If you intend to print or download the citation after viewing your results, press the **F9** key or the **spacebar** to mark the title as a desired citation before moving on to your next choice. If you change your mind, press either key again to unmark the choice.

To modify your search or enter a new search, press the **F3** key.

VIEWING SEARCH RESULTS

113

PRINTING OR DOWNLOADING

You may print as many citation records as you wish, and you may download up to 50 citation records at a time. You may also choose to simultaneously print and download your choices.

You may choose to print or download the short format (citation only) or the long format (citation and abstract). Follow these steps:

STEP ONE:

Prepare for Action

To print, insert your Copicard in the Copicard reader.

To download, remove the program disk and insert your MS-DOS formatted disk in the computer drive.

STEP TWO:

Choose the Citations to be Printed/Downloaded

If your are not planning to print/download all the citations found in your search, you must mark those that you want. Instructions for marking citations are given on the previous page, **VIEWING SEARCH RESULTS**.

STEP THREE:

Choose Your Options

Press the **F4** key. A window will appear with several options, including printing and downloading. A pointer on the left side of each option group shows the default choices. If you wish to choose a different option, use the ↑↓ arrow keys to move the highlight bar to your choice. In each option section, after moving the highlight bar to your choice, press the **F9** key; the pointer will move to your choice.

The last section of the menu is for downloading only. You may use the default name for your file, or you may type a name of your choosing (up to eight characters).

When everything is set the way you want it, press the ↵ key. If you have the pointer set on printing, printing will begin. If the pointer is set on downloading, the data transfer will take place.

To stop printing or downloading, press the ESC key. Printing or downloading will stop after a few seconds.

PRINTING OR DOWNLOADING

114

COMMAND SUMMARY

F1	**HELP**. Opens and closes HELP windows. HELP windows are specific to your current location in the program.
F2	**COMMANDS**. Press F2 repeatedly to see all of the active commands on the bottom line of the screen.
F3	**SEARCH**. Press F3 to begin a new search or modify the current search.
F4	**PRINT OR DOWNLOAD**. Allows you to print or download your search results.
F7	**CITATION WITH ABSTRACT**. View complete citations retrieved in search. This command skips the list of titles.
F10	**RESTART**. Clears all previous searches and returns to the Main Menu.
Shift-F3	**PREVIOUS SEARCH**. Allows you to return to a previous search exactly where it was terminated.
↵	**ENTER**. Completes the entry and moves forward through the program.
Esc	**GO BACK**. Moves backward through the program, retracing your steps.

QUITTING

Press **F10** to return to the **Main Menu**. Select **Exit**.

UNIVERSITY OF HOUSTON

CD-ROM databases

ABI/INFORM STRATEGY WORKSHEET

A workspace has been provided for you on the back of this form. The steps below will help you to plan your search strategy.

1. STATE YOUR TOPIC

Write down what you want to know. Example:

Do salary increases affect employee productivity?

2. IDENTIFY THE MAJOR CONCEPTS IN YOUR TOPIC

Mark the concepts in your statement. If you find that you have four or more concepts, discuss your topic with the staff at the Reference/Information Desk. Example:

Do salary increases affect employee productivity?

3. CHOOSE SEARCH TERMS FOR EACH CONCEPT

Use the Search Inform thesaurus, available near the workstations, to help you with selecting terms. List synonyms and other acceptable words for each concept. Some concepts will have many alternative words or phrases; others can be described using only one term. Example:

Concept 1	Concept 2
salary increases	productivity
raises	

When choosing terms, there is no need to choose a phrase that includes a word you have already chosen individually. For example, raises would retrieve merit raises.

Whenever possible, avoid choosing terms that include very common words, such as "management" or "marketing." These words, which appear frequently in business literature, can significantly increase the amount of time it takes to do your search. Less common words result in faster searching.

ABI/INFORM automatically searches singulars and plurals, regardless of how you enter a term. Some short words, such as "of, the, a," etc., are not searchable. They are known as stop words. Use the "pre/#" operator to allow space for stop words in any phrase you search. See the PROXIMITY OPERATORS portion of the ADVANCED SEARCHING section of the ABI/INFORM booklet or flipchart for an example of how to do it.

ABI/INFORM does not search phrases in random order. For example, in a search on "market share," you might also consider searching for the phrase "share of the market." You may search both word orders simultaneously by using the "w/#" operator. See the PROXIMITY OPERATORS portion of the **ADVANCED SEARCHING** section of the ABI/INFORM booklet or flipchart for an example of how to do it.

4. **FINE TUNE YOUR STRATEGY**

Read the **ADVANCED SEARCHING** section of the ABI/INFORM booklet or flipchart for instructions on how to truncate similar terms and how to limit your search to a specific field.

The **SEARCHING TIPS** section of the flipchart or booklet will provide additional suggestions for making your search efficient.

5. **SEARCHING ADVICE**

Use the connector OR to link the terms within each concept group. Use the connector AND to link different concepts.

We advise you to enter your concepts one at a time, then link the resulting set numbers by using AND. This will allow you to reuse concepts if the initial results are disappointing. Example:

[1] salary increases or raises
[2] productivity
[3] [1] and [2]

Always enter the concept that you think will result in the least number of citations first. If it results in a fairly small number (under 50), it will usually take you less time if you skip the additional concepts and look at the citations.

STATE YOUR TOPIC HERE: _____

_____

LIST YOUR CONCEPTS AND TERMS HERE:

CONCEPT 1		CONCEPT 2		CONCEPT 3
_____		_____		_____
or		or		or
_____	AND	_____	AND	_____
or		or		or
_____		_____		_____

Bill Jackson/October 1988
University of Houston Libraries

Self-Service Computerized Information Retrieval

New York University
Elmer Holmes Bobst Library

Information Bulletin 6d

What is computerized information retrieval?
Many print sources are now available in a computerized form, providing powerful and flexible searching capabilities for periodical indexes, numeric data and general reference sources. At Bobst Library, users have access to computer-based information services in two formats: 1) through a compact disk (CD-ROM) attached to a microcomputer, or 2) online via a telecommunications connection to a remote database. Searching may be done by a librarian [see Information Bulletins 6a - c] or, for some databases, by individual researchers.

What is CD-ROM?
With CD-ROM (Compact Disk-Read Only Memory) technology, data are digitally encoded onto a plastic disk and may be "read" or played many times but cannot be erased or changed by the user. If you've ever used a periodical index in a paper edition, you're already familiar with the content of CD-ROMs. To search a CD-ROM, one needs a CD player (different from those used for audio compact disks) and a microcomputer.

What are online information retrieval systems?
A search of an online database is achieved by connecting via telephone lines to a host computer which may be located anywhere in the world. The connection is generally a high speed transmission which allows many users simultaneous transactions with the database. Online databases are typically more up-to-date than CD-ROMs and cover many subjects and disciplines.

Products and Locations:

Reference on Main:	Dissertation Abstracts CD
Reference on 7:	PAIS CD, Government Publications Index CD, CIS Congressional Masterfile CD, Dow Jones News Retrieval Online, Business Connection Online
Reference on 9:	Medical Connection Online

See page 4 for descriptions of each product.

Availability

• All products are available during Reference Service hours:

Mon - Thurs	10 am - 9 pm
Fri	10 am - 6 pm
Sat	10 am - 6 pm
Sun	2 pm - 8 pm

• Sessions are limited to 30 minutes. *Be sure to allow time to both search the database and print your citations within the 30 minute session.*

Cost
Searching the products listed above is free to all qualified users. You may want a 5.25" floppy diskette [available at the Bobst Copy Center or the NYU Bookstore] to download. The diskette must be formatted before attempting to download.

Use
These products require no prior computer knowledge. There are several instructional items available, including: user guides, manuals and on-screen help options. The library staff will assist you with connecting to the various services. Other assistance may be available upon request.

Advantages of Computer Searching

Computerized database searching offers many advantages to the library researcher, among them speed, currency and convenience. The computerized database usually encompasses multi-year cumulations of a particular index, relieving the reseacher of checking individual annual indexes. It also allows searchers to combine key terms in order to find a unique subset of items, an approach frequently difficult to accomplish through searching a print source. Finally, the ability to download to disk means that you can easily incorporate the information into your own word processed documents.

Keyword Searches

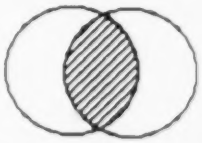
Over the years, every database has developed its own list of specialized (also called controlled) vocabulary to describe the subjects covered. When using the print indexes, you *must* use terms found in the authorized thesaurus for successful retrieval. While it's always preferable to use the thesaurus term, computer searching is much more flexible. It allows you to look for a keyword from the title or abstract and is not limited to controlled vocabulary.

Boolean Logic

By connecting keywords (or thesaurus terms) together, you can refine your search to retrieve only the most pertinent citations. Boolean Logic revolves around three terms: *"AND" "OR" "NOT"*

AND selects the intersection of two concepts:

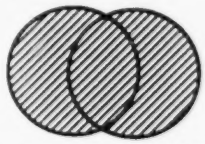
Japan **AND** computers



This finds articles that mention both Japan and computers somewhere in the database entry.

OR creates the union of several concepts:

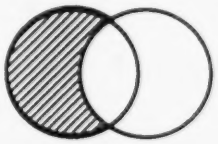
Japan **OR** Korea



This is used to establish one large set of similar or synonymous terms. Entries will mention either Japan or Korea or both.

NOT eliminates any entries that contain a mention of that concept:

China **NOT** Taiwan



Use **NOT** sparingly as potentially useful items can be deleted.

Truncation

When you search for a word on a computerized database, the computer "finds" any occurence of that word exactly as you have spelled it. It matches letter for letter. For words with the same root, [*television, televisions, televise, televised,* etc.] it is very tedious to type each word. ***Truncation*** allows you to match the letters you have entered and to complete the word ending. For example, by placing a truncation symbol (such as "**?**") after the root common to all of the words [**televis?**], the computer will retrieve all of the possible combinations listed above.

Each system has devised its own truncation symbol. Check each product's printed documentation for details. Be careful not to truncate the stem too much. A heavily truncated stem, such as **tele?** will pick up *telepathy, telegram, telephone* and other words extraneous to your search.

Refining or limiting techniques

Some databases allow you to restrict your search results to articles from a specific year, journal title, language, institution, or agency. See printed instructions for each database for details on modifying or limiting capabilities.

Basic Strategy for Computer Searching

Step 1. Analyze your topic.

Divide your topic into its main concepts, and be as specific as possible. Here's an example of a topic refined from a broad to a narrow statement:

Business boom in the Far East
Growing high-tech industries in the Far East
Competition for US firms from computer industries of Japan and Korea

Step 2. Select the appropriate computerized information service.

Each database covers a particular subject area. You need to choose the right one(s) for effective search results. Consult page 4 for details or ask for assistance at the Reference Desk.

Step 3. Select terms.

Consider the many different terms you could use to describe your topic. For help in identifying related terms, check the Library of Congress Subject Heading volumes at the Reference Desk or, if it exists, a thesaurus specially designed for the database you're using. Make a list of terms for each of your main concepts:

Far East	*technology*	*competition*
East Asia	*electronic data processing*	*barriers to entry*
Korea	*computer industry*	*cartels*
Japan	*computers*	*industrial concentration*

Step 4. Connect concepts with Boolean Logic.

(Japan or *Korea)* and *(computer industry* or *electronic data processing)* and *competition*

In this sample search, you can find information on competition in the computer industries of both Japan and Korea.

Step 5. Use a search command to initiate the search request.

Like everyting else, search commands are different for every system. Follow on-screen instructions or check the appropriate printed guide.

Step 6. Modify search results if necessary.

Many products allow you to change your search at any time during the process. For instance, you can limit by date or language or add and delete terms. Use a printed guide for help with this.

Step 7. Print or download your search results.

To print or download your search results, check the on-screen help options or printed guides for further instructions. *To download, you will need a DOS formatted 5.25" diskette.*

Step 8. Review search results.

Evaluate your search results to see if you want to search additional terms or other databases.

Step 9. Search BobCat and/or UNISEC to determine if NYU subscribes to the journals you found in your search.

For details on how to do this, see Information Bulletin #4.

Step 10. Consult relevant printed sources as necessary.

To locate information not available in a computerized form, ask for assistance at the Reference Desk.

Database Descriptions

Business Connection
This interactive online service provides information on both private and public companies and on products. Company information may include financial profiles, late-breaking news, and investment research reports. Product coverage may be market data, manufacturers, design information and more. Material provided by the service comes from some of the most widely-known business publishers including Moody's, Standard & Poor's, Dun & Bradstreet, Predicasts and others. The system is menu driven using company name, ticker symbol, product or industry code or description or certain types of keywords. Results may be actual financial/statistical date, full-text reports or citations to journal articles. Information is continuously added or updated as it becomes available.

CIS Congressional Masterfile
This historical archive on compact disk includes complete indexing and citations for almost 400,000 congressional hearings, reports, documents and committee prints issued by the U. S. Congress between 1789 and 1970. The disk may be searched by committee or subcommittee name, specific Congress, date subject and/or keyword(s). Microfiche is available for all documents cited on the disk.

Dissertation Abstracts Ondisc
A multi-disciplinary listing of doctoral dissertations completed at accredited North American colleges and universities. It covers dissertations from 1861 to the present. A summary, or abstract, is included for each dissertation beginning with the year 1980.

Dow Jones News Retrieval Service (DJNS)
DJNS is an interactive information retrieval system that consists of a variety of business information on publicly held U. S. companies. The system is completely menu driven and includes recent business and economic news, corporate earnings, dividends and other market figures, comparative corporate financial and industry data, extracts from 10-K reports, recent analyst reports on companies and current and historical (up to one year) stock quotes and indexes. The service is continuously updated as new information becomes available, but, in most cases, only the most current information remains in the database.

Government Publications Index (GPO)
The GPO CD contains citations from 1976 to the present for publications of the United States government including books, reports, studies, serials, maps and more. It is the electronic version of the *U.S. Monthly Catalog of Government Publications*, a comprehensive index to publications from all of the federal agencies and the Congress. The disk is updated bimonthly and can be searched by author, title, subject, agency, publication date or keyword (s).

Medical Connection
An online system to 30 selected databases, including, medicine (Medline 1966 to date), chemistry (CA Search 1967 to date), biology (Biosis 1969 to date), technology (Compendex 1970 to date), physics (INSPEC 1977 to date), nursing (Nursing/Allied Health 1983 to date), food science (Food Sci/Tech Abstract 1969 to date), and computer science (Microcomputer Index 1981 to date). All databases include abstracts with the exception of Medline which includes abstracts after 1975 only and CA Search which contains no abstracts.

PAIS
PAIS (Public Affairs Information Service) on CD-ROM includes citations to journal articles, books and government documents in subject areas related to public administration, public policy, foreign relations, urban studies and other public interest topics in the social sciences. The database can be searched by author, title, subject headings, journal name, publication date, type of document or keyword (s). Citations may include a very brief statement (usually only one sentence) indicating the main theme of the cited material. Coverage begins with 1972 and updates are quarterly. Substantial foreign language material is included.

Other products may be available on a trial basis.

Data Files currently available include:

1. 1990 Test File: Public Law 94-171 Test Census of Los Angeles County

2. 1990 TIGER/Line File – Geographic & Cartographic Information Data Base

3. 1980 Census of Population and Housing (Summary Tape Files and Microdata)

4. 1970 Census of Population and Housing (1st-5th Census and Microdata)

5. Migration Flows (1975-1980)

6. 1980 Equal Employment Opportunity

7. 1982 Economic Censuses

8. 1986 County Statistics File – data from various federal agencies

9. County Business Patterns (1976-1986)

10. County and City Data Book (1947-1983)

11. 1982 State and Metro Area Data Book

12. Current Population Survey (1980-1988)

13. Geographic Equivalency Files

14. 1980 Geographic Base Files-digitized DIME files

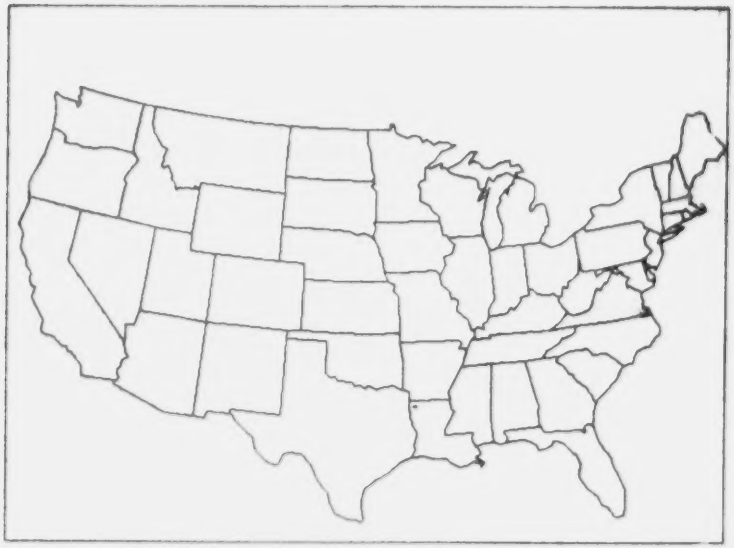
15. Master Area Reference Files

JZ:4/88;rev. 1/89

The Ohio State University
LIBRARIES

Guide to the

United States Census Data Center



OHIO STATE UNIVERSITY

Information Services Department
1st Floor, Main Library
1858 Neil Avenue Mall
Columbus, Ohio 43210
(614) 292-6175

124

Introduction

The University Libraries, through the Information Services Department in Main (Thompson) Library, is the University's designated Census Data Center for U.S. Census computer tapes on campus. Data files containing U.S. Census data are obtained and disseminated as a result of the Joint Statistical Agreement between the Ohio Department of Development's Ohio Data Users Center and The Ohio State University. Any Ohio State student, staff, or faculty member can request a search of the machine-readable files free of charge. Presently, nearly 200 files are maintained, archived, and ready for use. The Center receives all Ohio data and a large selection of national files.

LOCATION: Information Services Department
1st Floor, Main Library
1858 Neil Avenue Mall
Columbus, Ohio 43210
(614) 292-6175

The Center Provides:

- access to the data files
- technical documentation and codebooks identifying the file structures of the tapes
- technical assistance, consultation, and sample programs
- census tape processing and data analysis
- training in access to and in applications of census data
- seminars and classes on census data availability and retrieval
- workshops on data sources for selected topics, e.g., data on the elderly or data on women
- up-to-date inventory and catalog of available data files, computer software, maps and publications
- tape copying for dissemination to other universities throughout the state

130

Why Data on Tape?

When dealing with census information, data on tape is more feasible than printed reports for a number of reasons. First, printed reports may not carry the level of geographic detail desired. Additionally, they are not organized in a manner appropriate to the acquisition of information, e. g., data for the fifty states would require a search in fifty separate volumes. Data on tape avoids these problems.

Furthermore, certain forms of data are simply not available in print, e. g., zip code data, population centroids, and microdata. And, while certain data are available in print as well as on tape, the tape version provides more extensive information in such areas as graphics or computer mapping.

Data on tape makes possible more cross tabulations, e.g., income in relationship to gender. It also permits manipulation of the data on the computer—this saves time coding the data.

ACADEMIC DATA SERVICE

Users' Guide

Version 1.0
September 1, 1988

A joint publication of the Systems Development (SyD) Division of
Academic Information Resources (AIR) and Stanford University Libraries.

ACADEMIC DATA SERVICE USERS' GUIDE

TABLE OF CONTENTS

INTRODUCTION 1

FINDING DATA FILES AT THE STANFORD LIBRARIES
- Knowing About Specific Data Files
- Verifying the Availability of a Data File
- Data Files Not Among Current Library Holdings 2
  - When Your Data File Arrives 3

REQUESTING AN ADS DATA FILE
- The ADS Host Computer: WATSON
- The ADS Request Form
- Loan Periods 4
- Loan Renewals

USING YOUR ADS DATA FILES ON WATSON
- Learning How To Use WATSON
  - Introduction to VM/CMS
  - Consulting Services
  - ADS Disk Space Information 5
- Working in the WATSON VM/CMS Environment 6
  - Statistical Software
  - File Naming in VM/CMS
  - SAS And Your ADS Data File 7
  - SPSSX And Your ADS Data File
  - Scratch Files
  - LISTING Files 8
  - Temporary Disk Storage

USING ADS FILES ON OTHER COMPUTERS 9
- Transferring ADS Data Files
  - To Another Host Computer
  - To a Microcomputer on SUNet
  - To a Microcomputer not on SUNet
  - By Magnetic Tape 10

APPENDIX A: Sample ADS Data File Request Form

INTRODUCTION

The Academic Data Service (ADS), a joint endeavor of Stanford University Libraries (SUL) and Academic Information Resources (AIR), is a coordinated service program for the acquisition and use of research materials in machine readable form (Machine Readable Data Files, or MRDFs). These are distributed to Stanford by the Inter-university Consortium for Political and Social Research (ICPSR), the Roper Center, and other national and international archives, and include census records, election returns, public opinion polls, and statistics of all types.

This guide is intended to assist you, the faculty, students, and staff researchers at Stanford, in using specific data files of interest to you. It describes the steps you typically go through to use the service, from knowing about a desired data file, to using the data file on the ADS host computer.

FINDING DATA FILES AT THE STANFORD LIBRARIES

Knowing About Specific Data Files

Researchers affiliated with Stanford University and interested in accessing machine readable data files usually come to the University Libraries with a certain amount of information about the data file they want. Some individuals are very knowledgeable about the data files made available by data archives and approach reference staff with a specific citation for the data file they want to use at Stanford. While data files are not currently catalogued in either the card catalog or Socrates, the on-line catalog, there are records in both locations for the codebooks which document some of the data files acquired by SUL. Individuals who have located records for a specific codebook then approach reference staff or a bibliographer to find out what they should do next.

Other researchers have a very general idea of the kind of data they need. These individuals talk to the reference staff to determine if a data file that meets their needs exists in the Library's collection or elsewhere. You might know that the University Libraries have already acquired a data file you want to use (perhaps from a colleague who has used the file before). Contact reference staff or the bibliographer who works with your department or school to find out how to get access to the data file.

Verifying the Availability of a Data File

Stanford University Libraries has been acquiring data files for many years. The Data File Directory (DFD), (Z881.S7925 1987, at the General Reference Desk, Green Library; Jonsson Library of Government Documents; Cubberley Education Library) contains descriptions of the data files acquired through 1987. Reference librarians can assist you in searching the printed DFD. The DFD is also available on-line as a SPIRES data base on the Data Center computer; this data base contains descriptions of data files acquired after the printing of the 1987 paper edition of the DFD. A users guide to the DFD is available at the libraries listed above.

Most data files have accompanying documentation. These codebooks are cataloged in Socrates, the on-line catalog, and shelved in one of the Stanford University Libraries (most often Green Library). One copy of each codebook will circulate to individuals with valid borrowing privileges.

Once, with the assistance of Library staff, you have located the data file you desire, you can request that it be made available to you. See "Requesting a Data File", below, for instructions on making the request.

Data Files Not Among Current Library Holdings

There are other data files not currently in the Library's collection that may be obtained from various organizations. The Library's acquisition decisions are based on support for the graduate academic and research programs at Stanford, individual doctoral research support, cost of acquiring the data file, and the assumption that any data file purchased will become part of the Academic Data Service and be accessible to the Stanford community. Some sources of data files are described here.

Stanford is a member of the Inter-university Consortium for Political and Social Research (ICPSR), and Eddy Hogan, at the General Reference Desk in Green Library, is the Official Representative. This large data archive organization produces a yearly listing of its data file holdings, the Guide to Resources and Services (H61.G85, at the General Reference Desk, Green Library; Jonsson Library of Government Documents; Cubberley Education Library; and there is a circulating copy in Green Library stack).

Data files from the ICPSR may be requested directly through Eddy Hogan or Eric Heath of General Reference. Librarians in other locations may verify that the data files desired by their clients are listed in the ICPSR Guide and make order referrals to Eric.

Stanford is also a member of the Roper Center for Public Opinion Research and data files may be acquired through this center. Roper produces some documentation about its data file holdings, but nothing as developed as the ICPSR Guide. If you are interested in a study conducted by the Roper Center, contact Eddy Hogan or Eric Heath to verify the availability and cost of the data file.

Generally, ICPSR data files are inexpensively purchased while data files from Roper usually cost more than $100 each. Roper will also produce reports especially tailored for the individual researcher, thus eliminating the need for the researcher to process data. These reports are very expensive and are approved only after serious consultation, negotiation with Roper, and written support by a Stanford faculty member. See Eddy Hogan for more information.

Often agencies of the United States Federal government have produced data files that are not made available through ICPSR or the Roper Center. See Joan Loftus in the Jonsson Library of Government Documents.

International governmental agencies, such as the World Bank, may also produce data files not made available through ICPSR or Roper. See David Rozcuszka in the Jonsson Library of Government Documents.

You may know of more elusive data files that are not included in the categories mentioned above. Again, consult first with Eddy Hogan. The General Reference Department has a collection of data archive catalogs which may be helpful in locating the data file you need.

Sometimes a scholar or group of scholars have produced a data file as part of a research project sometime in the past. No attempt will be made by Library staff to locate such scholars. As a general principle, data files produced by individual scholars that are not well documented and not formatted according to specifications acceptable to the Academic Data Service will not be acquired.

When Your Data File Arrives

After your data files are received by the Stanford University Libraries, you will be sent a note stating that the data files have been sent for processing into the ADS library. When these have been processed, an entry for the new data files are made in the DFD, providing you with the information needed to request the files be made available for personal use. (the following section, "Requesting a Data File," provides instructions for making this request.) The codebook for the data file will be rush cataloged and the requestor will notified when it is becomes available at the Loan Desk of the library where the book will reside.

REQUESTING AN ADS DATA FILE

The ADS Host Computer: WATSON

Computer access to ADS machine readable data files is provided through the ADS host computer, WATSON, an IBM 4381 computer running the VM/CMS operating system. When you request a data file, you are provided an account on WATSON and the data file is loaded onto disk for you there.

The Academic Data Service may be used only by Stanford University faculty members, students and staff. The WATSON computer may not be used for research that is sponsored (funded) by any agency external to Stanford (i.e., government or industry sponsored research). WATSON accounts are granted to individuals. You may use only your own account. You may not grant permission to any other person to use your account.

The ADS Request Form

Once you have located the desired data file in the Data File Directory, you need to request that it be loaded onto disk and an ADS WATSON account be created for your use. You must know the following three pieces of information from the DFD entry for each file you request: (1)the File Title, (2)the Tape Number, and (3)the File Number.

You request a data file by completing and handing in an Academic Data Service Data File Request Form. This form is available at the Reference Desk of Green Library, the Jonsson Library of Government Documents, and the Cubberly Education Library. See Appendix A for a sample filled-out form. Alternatively, if you already have a WATSON account, you may consider initiating the request electronically by copying a Request Form Template, modifying this template file to include the request information, and sending it to the WATSON account ADS. To do this, type

COPYFILE *ADS REQUEST P ADS REQUEST A*
XEDIT *ADS REQUEST A*

SENDFILE *ADS REQUEST A* **TO** *ADS*

Requests are processed daily, Monday-Friday, and an account is generally created within one working day. Data files may take longer depending on the availability of disk space.

PLEASE NOTE: Although you may already have a WATSON account as an affiliate of the departments of Communications, Economics, Political Science, Sociology, the Law School or IMSSS, a separate account will be created for ADS service. The requirement of an account specifically for ADS use is driven by the large amount of disk space typically needed for

working with the characteristically large data files, and to enable ADS to record computer usage rates and patterns.

Loan Periods

ADS data files are library holdings, and as such, are provided to you on a loan basis. You are given access to the data files and WATSON for specified "loan periods". The loan period for a data file is five weeks; the loan period for your WATSON account is ten weeks. These dates are displayed on your terminal when you log in. The assumption is that the early stages of processing a raw data file include the generation of files containing the data in forms used for all subsequent processing (e.g., data file samples or subsets, SAS data sets, SPSS^{x} system files); such files derived from the original data are stored on the private disk of your WATSON account and remain available for the entire account loan period.

Loan Renewals

You may request one week renewals toward the end of the five week data file loan period, or the ten week period of your WATSON account by sending electronic mail to account ADS or by sending an ADS Request Form through the Library. Such requests are considered on a case-by-case basis, and approval depends on the availability of ADS resources on WATSON.

USING YOUR ADS DATA FILES ON WATSON

Learning How To Use WATSON

WATSON is host to a number of current statistical packages and programming languages, including SAS, SPSS^{x} , BMDP, Minitab, VSFortran, Pascal/VS and PL/1. Terminals for access to WATSON are located on the 2nd Floor of Sweet Hall. Other terminals and terminal emulators connected to the Stanford network, or those with modems, can be used to work on WATSON.

Introduction to VM/CMS

The AIR documents entitled, "Introduction to VM/CMS Services", and, "VM/CMS - WATSON Connections" (available on the 3rd Floor of Sweet Hall) provide background information on subjects such as connecting to WATSON from a terminal or dial-in, and use of the text editor, XEDIT. Once you are logged onto WATSON, an extensive on-line help facility is available for your use. To see an overview of how the help system works, use the command:

HELP *INTRO*

Another good place to begin learning about WATSON is by using the on-line news system. Use the command:

NEWS

Classes on VM/CMS use are scheduled on a quarterly basis. To see the current schedule, use the NEWS command and select the Intro Classes topic. A printed schedule is also available on the 2nd Floor of Sweet Hall.

The WATSON text editor, XEDIT is used to create and save files. XEDIT instruction is available through an on-line tutorial. To access this tutorial use the command:

Consulting Services

Consultants are also available to help you with general and specific questions regarding the use of WATSON and its operating system, VM/CMS. To find out who these consultants are and their schedules, use the command:

HELP *CONSULT*

Statistical package consultants are also available to help you with questions regarding statistical software such as SAS and SPSS^{X} ; call 723-1055 or visit Sweet Hall, 3rd Floor.

For assistance with an ADS WATSON account or an ADS data file, contact Felix Limcaoco, 3rd Floor, Sweet Hall, 723-4942 (electronic mail: ADS@WATSON).

ADS Disk Space Information

There are three storage areas that you should be familiar with while using your ADS WATSON account: your default minidisk A (A-disk), the ADS data disk(s) which stores your requested data files (C-disk), and temporary disk space (T-disk).

A-Disk

Your A-disk allows you unrestricted READ and WRITE privileges and is created at the same time as your ADS account. You should be able to use the A-disk to store all the working files, data extractions, statistical package system files, listings, logs, and programs, that you create while working with your data file. This disk remains available as long as your account is active.

C-Disk

Your requested data file is loaded onto a separate data disk, a C-disk. You have READ ONLY privileges for the files on the C-disk. READ ONLY privileges imply that you cannot write to the disk, preventing you from corrupting the data file in any way with editors or programs. Depending on the files you request, you may be given the READ ONLY link to more than one disk, and these additional disks will have filemode letters other than "C". In addition to your account and disk access due dates, the contents of all of your ADS READ ONLY disks are also summarized when you log in. These disks remain available for 5 weeks. Directing the software you use to read data in from your data disks and to route output to your A-disk will be discussed in a later section.

T-Disks

Although space provided by your A-disk should generally suffice, you may need to access temporary disk storage for short periods of time. More information on setting up and using T-Disks will follow later in this document. If you find that you need to access a temporary disk please remember to release the disk space at the end of your CMS session. BE CAREFUL: As the name implies, files on the T-disk are erased at the end of your session. Copy any files needed later to a permanent READ/WRITE disk, like your A-disk, before logging off.

Working in the WATSON VM/CMS Environment

Statistical Software

To see a list of the software available on WATSON use the command:

HELP *PRODUCTS*

On-line help for SPSS^X , SAS, BMDP, and Minitab are found by using the command:

HELP

and selecting the desired topic.

Programs such as SAS and SPSS^X can be run in both interactive and batch mode. BATCH processing can be of particular use to the ADS user. When working with long programs or extremely large ADS data files, batch processing may be the only reasonable method for processing your programs. For more information on batch processing use the command:

HELP *BATCH*

An example of using batch processing on SAS and SPSS^X jobs is also given and discussed within the appropriate on-line help file. Use the commands

HELP *SAS*

or

HELP *SPSS^X*

to view these help files.

Further information on the use of the statistical packages in the VM/CMS environment can be found in the following AIR documents: "Using SPSS^X on the VM/CMS System", "Using BMDP on the VM/CMS System", "Using SAS on the VM/CMS System", and "Using MINITAB on the VM/CMS System." These documents also point you to consultants and other sources of more detailed information.

File Naming in VM/CMS

In the VM/CMS operating system each file is named using a three part name:

filename filetype filemode

These specifications are often abbreviated *fn ft fm*. You give all files an arbitrary filename, however the filetype and filemode names may have special significance. The filetype can name the application or type of file. Some applications require specific filetypes such as: SAS, SPSS^X or EXEC. The filemode names the minidisk where the file is stored. The ADS data file is normally stored on a disk accessed to your CMS session as the C-disk, so you will specify C for the filemode when naming these files. If other data disks are linked and accessed by your account, use the appropriate filemode letter to access the data files on them. All files that you create, like programs or system files, are generally stored on the disk

accessed as the A-disk so the filemode is A. If the filemode is not specified on the command line the default, A-disk, is used.

SAS And Your ADS Data File

To access an ADS data file which has been loaded into your C-disk, you need to issue a FILEDEF command within the SAS program. The CMS FILEDEF is used to link the SAS System through the operating system to the location of permanently stored files. The syntax of the FILEDEF statement as used in a SAS program is:

FILEDEF *DDname* **DISK** *fn ft fm*

In constructing the FILEDEF statement you will use a *DDname*, an arbitrary name for raw data files (or a LIBREF when referring to SAS system files) and the *fn ft fm* of the file to be used. For example, if a raw data file on your ADS C-disk is named *SURVEY84 DATA* , the *DDname* can be arbitrarily chosen to be *rawdata* and the FILEDEF statement in your SAS program would be:

FILEDEF *rawdata* **DISK** *SURVEY DATA C*

You are then able to access this file through the standard INFILE statement by pointing to the DDname *rawdata* from the FILEDEF.

INFILE *rawdata*

SPSS^X And Your ADS Data File

To access an ADS data file which has been loaded into your C-disk, you need to use the SPSS^X FILE HANDLE command within your SPSS^X program. This command is used to define the location and the name of the file to be read or written by the SPSS^X program. To use the FILE HANDLE statement an arbitrary *nickname*, or *handle* is used. The syntax of the FILE HANDLE statement as used in a SPSS^X program is:

FILE HANDLE *nickname* / **NAME** = *'fn ft fm'*

The file is defined by specifying the *fn ft fm* of the file to be used. For example, if your ADS data file on your C-disk is named *SURVEY84 DATA* and the *handle* is arbitrarily chosen to be *rawdata*, the corresponding FILE HANDLE statement in your SPSS^X program would be:

FILE HANDLE *rawdata* / **NAME** = *'SURVEY84 DATA C'*

You are then able to access this file through the standard DATA LIST statement by pointing to the handle *rawdata* from the FILE HANDLE:

DATA LIST FILE = *rawdata*

Scratch Files

Both SAS and SPSS^X create and use temporary system files during execution. These files are normally written to your A-disk and are erased after completion of a program. They are generally never seen by you; however, if the program is unable to write these temporary files to the disk because of insufficient room on the A-disk, an error message similar to the following will appear:

DISK 'A(0191)' IS FULL

This means you would then need to make room on your A-disk by using the ERASE command to delete files no longer needed, or acquire a temporary disk (T-disk) for the temporary storage of these files.

LISTING Files

If you submit a noninteractive SAS program residing on your A-disk to SAS, after execution the A-disk will contain the SASLOG and LISTING files under the filename of your SAS program. Similarly, SPSS^X will place the listing file on the same disk as the program file, and will use the same filename as that of the command file with the filetype of LISTING. You should be aware that these LISTING files can be quite large if you request many procedures within a single run. If they are not required on-line it is suggested that you print the files and erase them from your A-disk.

If SAS or SPSS^X tries to write a LISTING file to your A-disk and the disk is full, you will receive a message on your terminal screen.

Temporary Disk Storage

If you are unable to write either system files or output files to your A-disk, and you have erased any files that are no longer needed, you may want to use a temporary minidisk. The temporary disk is a valuable resource allowing you to expand the available minidisk space. For more information on how to obtain a T-disk use the command:

HELP *TDISK*

After you have obtained a T-disk you need to be aware that when you LOGOFF the CMS session the DISK and ALL of its contents are erased. If you want to store any of the files for future use, you can copy them to another permanent disk or to tape.

Perhaps the easiest way to utilize the T-disk is to follow these steps:

1. Copy your program file onto the T-disk using the COPYFILE command:

   COPYFILE *fn ft* A *fn ft* T

2. Rename the program file left on the A-disk using the RENAME command:

   RENAME *fn ft* A *newname newtype* A

3. Execute the program named on the T-disk.
4. Look for the output files generated by the program on the T-disk.

Please note that if you do not RENAME the original program file still on the A-disk as suggested in step 2, when you run the program, you will run the copy of the program file on the A-disk, not the T-disk copy where all additional available space is located.

Programs that are copied to the T-disk can still point to files on the A- or C-disks using either the FILEDEF or FILE HANDLE statements. The location of the file to be read or written must be coded on these statements using the appropriate filemode.

Remember that after the completion of the CMS session the T-disk and all its contents are erased.

USING ADS FILES ON OTHER COMPUTERS

Even if you indicate in your request form that you will be using another computer for your work on ADS files, you will be given a WATSON account. Since you will not be processing data files on WATSON, however, a smaller A-disk will be assigned to you. Your account and disk links will be active long enough for the file transfers to be made, expiring 2 weeks after your account is opened.

You are encouraged to use WATSON's up-to-date software facilities for analyzing your ADS data. If you prefer to use ADS files on another computer, you will have to transfer your files from the WATSON account to it. The following section provides information on how this can be done.

Transferring ADS Data Files

Please note the files requested through the ADS are typically quite large; before moving the data to another computer, be sure it can handle the storage demands.

To another host computer

The File Transfer Protocol (FTP) is one method to electronically transfer files from WATSON to another SUNet host (e.g., Lear, Othello, How) which supports the same protocol. Type:

HELP *FTP*

to obtain a description of the procedures and options for using this program.

To a microcomputer on SUNet

FTP may be used to transfer data from WATSON to a microcomputer which is connected to SUNet. The programs, SU-PC/IP and SU-MAC/IP (available from Networking, Pine Hall) are used on the IBM PC and Apple Macintosh respectively. Each program comes with a guide to its use.

To a microcomputer that is not on SUNet

The file transfer protocol, KERMIT, which is supported by WATSON, can be used for copying files to and from microcomputer disks. Documentation for this software can be found by typing:

HELP *KERMIT*

Another CMS command,

KERMIT DOC

lets you browse the KERMIT documentation in the XEDIT environment.

By Magnetic Tape

It may be more convenient or even necessary to transfer the data file to another host via magnetic tape. The WATSON tape drives are self-service, and are located in the tape drive room in the CERAS building. You must supply your own tapes; they are available from most stores that sell data processing supplies. To find more information on creating tapes in the VM/CMS environment use the command:

HELP *TAPES*

When creating copies of any data file on tape be sure that the tape you create is readable by the alternate host machine. The tape facilities and tools on WATSON allow you to create tapes at either 1600 or 6250 BPI, and in either CMS, unlabeled, or IBM Standard Label format. Check with the systems staff of the alternate host machine to learn the tape formats supported there.

APPENDIX A: Sample ADS Request Form

ACADEMIC DATA SERVICE
DATA FILE REQUEST

Please leave completed forms at the Reference Desk

Name: *Ronald Research* **Date:** *10/10/88*

Stanford Department: *Stanford Center For Chicano Studies*

Check One:

[ ] Faculty	[x] Doctoral	[ ] Masters	[ ] Post Doc
[ ] Undergrad	[ ] Visiting Scholar	[ ] Staff	

Phone Number(s): *home ( ) ###-#### work ( ) ###-####*

A new account on the WATSON computer will be created for you exclusively for use with your data files. If you have an existing WATSON account, please choose a different Account Name for this.

Account Name: *RONRES*
(not more than 8 alphanumeric characters, no spaces, the first character must be a letter.)

Password: *SOMEPASS*
(between 4 and 8 characters long; any characters are allowed, make sure that you remember it.)

You can try loggin on to WATSON using the Account Name and Password you've supplied in 1-2 working days. A message will be displayed when you log on telling you how to access your data files, when your link to these files expires, and when your account expires. If you are unable to log on please contact a WATSON consultant.

Will you be using your ADS files on WATSON or transferring them to another computer?

[x] WATSON [ ] Other

Data File Request:
Obtain the information requested below from the Data File Directory entry for the file. Ask the Reference Desk for help if needed.

FILE TITLE	TAPE NO.	FILE NO.
Current Population Survey: 1980	*AS2506*	*1*
Current Population Survey: 1983	*AS2879*	*1*

FOR OFFICE USE ONLY

Created:_____ C Size:_____
C Expire:_____ C Number:_____
Acct Expire:_____ Renewal:_____

138

User's guide to the University of Toronto
Data Library Service for EPAS users.

Laine G.M. Ruus
draft ed. 2/89

A. Introduction

The University of Toronto Data Library Service (DLS) was established in the summer of 1988. Its collection is small but growing!

B. Available data files

A brief title listing of data files in the DLS collections is available on EPAS.

1. To access the DATAFILE LIST

In order to access the listing, log on to EPAS and enter:

GETME DATALIB

This will give you access to the DLS's public minidisk, and you can then look at the list of available data files by entering to CMS:

X DATAFILE LIST *

If you intend to use the files on the DLS's 192 disk on a regular basis, you may wish to add GETME DATALIB to your PROFILE EXEC file. You can then enter:

QUERY DISK

to determine at what filemode the DATALB minidisk is attached to your user id, and use FILEL as usual to examine other files on that disk, such as the file containing this document, DATALIB MANUAL.

2. Contents of DATAFILE LIST

The data files in the DLS collection are listed in DATAFILE LIST in alphabetical order by title. Each title is followed by:
(a) a list of the physical files that make up the data file, indicating for each physical file whether it is:
- a data file (i.e. a flat character file), or
- an SPSS.x or SAS system file, or
- an OSIRIS dictionary (can be read by SPSS.x or SAS),

- an OSIRIS codebook (can be read by SPSS.x or SAS), or
- a codebook list file (usually including printer control characters), or
- some other type of file.

(b) for each physical file:
- the external and internal labels of the tape on which it is to be found
- the file sequence number of each file on the tape
- the datasetname of each file
- the physical format of each file, expressed as 'block type(blocksize,physical record length)'.

Take, for example, the following listing:

Census of Canada, 1981 - Public use sample tapes
- individual file
BN0149/CST001 4 CND.INDPUST.81 FB(24720,206)
- SPSS.x system file
BN0190/SPX001 4 CND.IND.81.SPX FB(20480,1024)
Codebook available in DLS room 4056, Robarts

This means that the Individual PUST from the 1981 census of Canada
- is available on the tape 'BN0149' (you will probably never need to know the internal volume label 'CST001')
- it is the 4th file on that tape
- it has a datasetname CND.INDPUST.81
- it is recorded on the tape in a fixed block format, with records that are 206 characters long, gathered into blocks that are 24720 characters long,
- similarly, there is an SPSS.x system file version of the same data on tape BN0190, with somewhat different characteristics,
- and a codebook describing the data file, including locations of the variables in the raw data file, is to be found in the DLS.

Each data file listed has associated with it a codebook which describes in detail the contents of the file and the record layout of the data.

In most cases, there will be two copies of the codebook in the DLS, room 4056, Robarts Library. One copy of the codebook may be borrowed - the other copy is anchored in the DLS (i.e. does NOT EVER circulate).

C. Using data files with statistical packages

Many files in the DLS collection are quite large; on the other hand, most users require only a specific subset of any particular file. Thus it is more efficient to use a statistical package to access and extract the subset of variables and/or cases required and transfer only the subset to your own disk space, than to copy the entire file into disk space.

Before you can efficiently use any data file, you must consult the codebook, directories, etc. which describe the content and layout of the data file.

The following section is intended to illustrate the process of extracting a subset with SPSS.x and SAS. For further information, users should refer to the relevant SPSS.x or SAS user manuals, and, where applicable, relevant CMS documentation.

1. Accessing tape files using SPSS.x

To access a tape file using SPSS.x, you should first consult the codebook to determine which variables it is that you need. Then, create two files on your 'A' disk:

In a file fn EXEC A, you should have a filedef statement for each tape file you wish to refer to, followed by an 'spssx' run command refering to a second file fn SPSSX A, in which are SPSS.x control commands.

For example, in order to access the raw data file of the 1981 individual PUST on tape, and save a subset of the file in your own disk space, you might create two files as follows:

Contents of the file SPXFREQ EXEC A:

```
/* */
'filedef in tapi sl 4 (blksize 24720 lrecl 206'
'spssx ind81'
```

Contents of the file IND81 SPSSX A:

```
title Census of Canada, 1981 - Individual PUST
file handle out / name='data spssxfil a'
data list file=in records=1 /
  prov 1-2
  cma 3-5
  sex 37
  marstat 38
  :
  other SPSS.x control commands
  :
select if (prov eq 59)
save outfile=out / map
finish
```

Finally, in order to run this, you should
(a) mount the appropriate tape (BN0149)(see subsection 3. below), and, once you have received back a message from the operator that the tape IS mounted,
(b) issue the command:

```
SPXFREQ
```


2. Accessing tape files using SAS

SAS allows you to accomplish the above in one file rather than two, as long as that file has ft SAS. For example, to read a tape file using SAS, and produce frequencies, the command file FREQ SAS A should contain:

```
cms filedef pust81 tapi sl 4 (lrecl 206;
data subset.ind81;
  infile pust81;
  input prov 1-2 cma 3-5 sex 37 marstat 38;
  if prov=59;
  :
  other SAS control commands..
  :
```

Once you have the commands in a file, you can then mount the relevant tape (see section 3. below), and once the operator has messaged you that the tape is mounted, enter the command:

```
SAS FREQ
```

3. Mounting a tape

(a) Before mounting a tape, you should make sure that the one tape drive available to EPAS is not already in use: issue the command

```
TELL EPASINFO QUERY TAPE
```

(b) If the tape drive is free, the system will respond

```
TAPE ACTIVE NOT FOUND
```

and you can enter the actual tape mount request, specifying the external label of the tape you wish to have mounted (as listed in DATAFILE LIST):

```
MOUNT FOR <your userid> <external label> 181 WITHOUT RING

e.g. MOUNT FOR LAINE BN0149 181 WITHOUT RING
```

(c) Once the tape has been mounted, you will receive a message TAPE 181 ATT, i.e. the tape has been attached on I/O unit 181 and is ready for use; having a tape mounted may take a little time. Once the tape is mounted, you should refer to it in filedefs as 'tapi'.

For further information on the MOUNT command, enter HELP MOUNT.

Be aware that you may only mount a tape between the hours of 9:00 am and 5:00 pm, Mondays through Fridays, and only when

no one else has a tape mounted on EPAS.

4. Rewinding a tape

Should your job be unsuccessful, for whatever reason, you cannot merely run the job again, as the tape will have stopped past the file you wished to read. You must therefore rewind it before attempting to have a job read the file again.

To do so, enter

```
TAPE REW
```

5. Detaching a tape

Be aware that your computer account is charged for each second that you have a tape mounted. Also, that on EPAS, access to tape drives is a very limited resource, and you may be preventing another user from having a tape mounted. Therefore you should detach the tape drive as soon as you are finished your job, and have made sure that it ran successfully. To do so, enter to CMS:

```
DETACH 181
```

D. Special cases

1. CANSIM University Base

The approximately 24000 time series of Canadian socio-economic data in the CANSIM University Base are stored in disk files on the EPAS system. Before accessing these data, you should have the following information from the CANSIM University Base directory:
- a list of the databank numbers (unique series numbers) you require,
- for each series, the start date of the data you require (and the start date of the series),
- and for each series, the periodicity of the series (i.e. whether monthly, quarterly, or annual).

Then, in order to access the data, to CMS, enter the command:

```
CANSIM
```

This will run a program which will compile a job to search the files for the time series you have requested.

The system will prompt you for:
(a) a jobname, which for the sake of illustration I will call JOBNAME (any name, eight characters or less, with no embedded blanks),
(b) an output format (i.e. whether you will be using TSP,

SAS, etc. for your analysis),
(c) a beginning year at which time series should start (enter the last 2 digits of the year only), and
(d) the databank (series) numbers you require (enter these one at a time).

For each series, the system will prompt you for:
(e) a start year,
(f) whether you wish the data converted to annual or quarterly data points, and
(g) an alternate name for the series (e.g. if you wish, you may rename series 'D464000' as 'CPI').

Finally, the system will compile a job to retrieve the data and put the commands in the file (under your id) JOBNAME EXEC and return you to CMS. To run the job, you simply enter:

```
JOBNAME
```

The time series retrieved, as well as file definition commands appropriate to the statistical package you will be using, will be written to the file JOBNAME OUTPUT under your id. You should Xedit this file to remove the system control lines as indicated in the file before using it as input to a statistical package.

Copies of the CANSIM university base series directory (latest) are to be found in the Government Documents Division (reference collection, CA1 BS C17) and in the DLS, both in the Robarts Library.

2. International Monetary Fund databases

At present, only the International Financial Statistics database is available at U of Toronto.

In preparation for retrieving data from this database, you should have prepared the following:
- a list of the countries/regions and an IMF 3-digit country/region codes for which you require data,
- a list of the alphanumeric subject codes for which you require data,
- a note of the periodicity and time span for which you require data (note: generally, annual time series begin in 1948, quarterly and monthly time series in about 1965).

The relevant codes can be obtained either from the printed publication (International financial statistics) or on-line. Indices of available time series are available in disk files on EPAS. Simply enter IMFCATS to receive a message as to the filemode of the appropriate minidisk:

Use 'FILEL * * fm' to access IMF catalogues

142

The time series available for each country/region are listed in separate disk files identified by a filename consisting of the 3-digit country/region code and a filetype consisting of the date of the latest version received; e.g. the disk file containing the time series available for Canada as of the January 1989 version of IFS has fn and ft 156 CAT JAN89. Each record includes an individual alphanumeric subject code (in columns 1 through 5), revision and source codes, the title of the time series, the country name, and the units in which the time series is expressed.

In order to extract time series from the database, which is stored on tape, enter:

IMFDATA

You will be in an interactive environment in which the system will prompt you for

(a) IMF 3-digit country codes. These should be entered one at a time, after each prompt. After each country code, you will be prompted for an alias. This is your chance to give the time series a meaningful country mnemonic that you will remember (e.g. country code '111', alias 'USA'). When you have finished entering country codes, enter '?' at the next country code prompt.

(b) IMF alphanumeric subject codes (as they are listed in the published IFS) and an alias for each. Enter these one per prompt, as for the country codes above. E.g. subject code '99a', alias 'GNP'. Note that you will receive all available subjects requested for all countries requested. The number of countries multiplied by the number of series must not exceed 700 (e.g. 10 series for 15 countries = 150).

(c) whether you wish monthly, quarterly, and/or annual data. Respond with a 'Y' (yes) or 'N' (no) to each of these prompts.

The system will create two files in your disk space, labelled COUNTRY CODES A and SUBJECT CODES A; it will also set up a batch job queued to run at a later time. DO NOT delete the two CODES files until after your job has run successfully. Normally, retrieval will take place overnight, and the output files will be available in your reader list the following morning.

Each subject code requested, at any one periodicity, in your retrieval will have been written to a separate file, but all countries/regions requested will be in that file. Thus if you had requested annual, quarterly, and monthly data for two subject series for four countries, you would have six output files, one file each of monthly, quarterly and annual data for each of the two subject series. Each filename consists of the series code plus an abbreviated periodicity

code; filetype is '1' - thus the file '99AQ 1 A' contains quarterly GNP data for the requested country/region codes. (Each output file can only hold time series for seven countries. If more than one output file is required, the filemode will be incremented e.g. '99AQ 2 A' holds GNP data for the next seven countries, etc.) The first line in each file echoes the series code, title, country codes, and unit of measure.

To retrieve the files, enter:

RL

and then RECEIVE the files as usual.

Printed monthly issues of the IFS publication, listing both three digit country and alphanumeric subject codes, are to be found in the Government Documents Division (UN9 HF 152) in the Robarts Library (5th floor).

E. Other services

1. Online database vendors

In order to provide access to computer-readable data in databases not held directly at U of T, the DLS has an account with I. P. Sharp, a remote online database vendor specializing in quantitative databases. Thus, data from the IFS Canadian, U.S., and international time-series databases available from this vendor can be retrieved for analysis on microcomputers or uploaded for analysis on a mainframe.

Charges for this service will be the same as for computer-aided bibliographic searches done by the UT Library, i.e. the user will be charged the full cost of the retrieval, plus a surcharge of $1.50. Should you wish the data downloaded to a microcomputer diskette (MS-DOS), you should bring a 5.25" ds,dd diskette (formatted) with you.

2. Acquisition of new data

The DLS collection is small, as yet. New acquisitions will partially be driven by user demand. The DLS has access to the collections of a number of major data file suppliers, including the ICPSR, Roper Center, Statistics Canada, National Archives of Canada, as well as those of many data archives and data libraries in Canada, the United States, Europe, Australia, etc. Should you require data not presently in the DLS collection, please contact Laine Ruus (tel. 978-5589, or via e-mail as LAINE@UTORONTO) to arrange to have the data acquired. Be aware that data acquisition can take some time, and should therefore be planned ahead of time.

3. Classroom use

Most significant data files are very large, and for classroom teaching purposes, subsets of a data file consisting of a selection of the original variables and a random sample of cases with a minimum of missing data, already predefined for use with a statistical package, and a set of groomed documentation that can be provided to each student is often a preferable alternative. The DLS will, given sufficient notice, be happy to prepare subsets of system files tailored to specific teaching needs.

4. Depositing data files

New data files, created in the course of research or teaching, may be of value to others for secondary analysis. If research based on the analysis of original data is published, it should be a normal part of the peer-review process that others have access to the data. Such data should be deposited with the DLS. The DLS will develop documentation as needed (and time allows), ensure the long term preservation of the data, as well as handle dissemination of copies to such other users as you indicate in the DLS deposit agreement.

5. Citing data files

Data files used in research, as any other source material in traditional published forms, should be cited in the bibliography of the work. Standard citation formats for data files have been established, and should be used. Please contact the DLS for further details.

A sample citation might look like:

Canada. Statistics Canada./ CANSIM university base [machine-readable data file]. Ottawa, Ont.: Statistics Canada [producer and distributor], Sept. 1988.
1 data file (ca 24,000 logical records) & supplementary documentation.

F. Location of the DLS

The DLS is located in room 4056 (4th floor) of the Robarts Library, 130 St. George Street.

Office hours: (approx.) 9:30 am - 5:00 pm

Telephone: (416) 978-5589

E-mail: LAINE@UTORONTO

144

SELECTED READINGS

Selected Resources

Andre, Pamela Q. J. "Optical Disc Applications in Libraries," *Library Trends*, 37, 3 (Winter 1989) 326-342.

Association of Research Libraries. *Microcomputer Software Policies in ARL Libraries*. (SPEC Kit No. 123) Washington, D.C.: ARL, 1986.

Chiang, Katherine B. "Computer Accessible Material in the Academic Library: Avoiding the Kludge," in *Energies for Transition, Proceedings for the 4th National Conference of ACRL* (Chicago: ALA, 1986) 67-69.

Dodd, Sue A. *Cataloging Machine-Readable Data Files*. Chicago: ALA, 1982.

Dudley, Claire S. "Microcomputer Software Collection Development," *CHOICE* 23, 5 (1986) 704-705.

Gerken, Ann. "What is a Data Archive and What Should the Information Specialist Know About Managing Locally Maintained Numeric Data Files?" *Database*, 11, 4 (August 1988) 60-65.

Heim, Kathleen M. "Social Scientific Information Needs for Numeric Data: The Evolution of the International Data Archive Infrastructure," *Collection Management*, 9, 1 (Spring 1987) 1-53.

Jones, Ray, and Kinney, Thomas. "Government Information in Machine-Readable Date Files: Implications for Libraries and Librarians," *Government Publications Review*, 15 (1988) 25-32.

Turner, Judith Axler. "Scholars Weigh Library's Role in Collecting Computerized Research Data," *The Chronicle of Higher Education*, 32 (July 16, 1986) 34.

Mallison, John C. "On the Preservation of Human- and Machine-Readable Records," *Information Technology and Libraries*, 7, 1 (March 1988), 19-23.

Mallison, John C. "Preserving Machine-Readable Archival Records for the Milenia," *Archivaria*, 22 (Summer 1986) 147-152.

Morris, Dilys E. "Electronic Information and Technology: Impact and Potential for Academic Libraries," *College & Research Libraries*, 50, 1 (January 1989) 56-64.

Neff, Raymond K. "Merging Libraries and Computer Centers: Manifest Destiny or Manifestly Deranged? An Academic Services Perspective." *EDUCOM Bulletin*, 20, 4 (1985) 8-12, 16.

Paskoff, Beth M. "Microcomputer Software in Library Collections," *Library Trends*, 37, 3 (Winter 1989) 302-315.

Research Libraries Group. *Managing a New Library Resource: Results of the RLG Machine-Readable Data File Project in Six Member Libraries.* Mountain View, CA: RLG, 1989.

Talab, Rosemary. "Back-ups: A Controversial but Necessary Part of Software Collections," *Small Computers in Libraries,* 7, 3 (March 1987) 36-39.